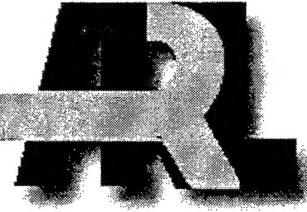ARMY RESEARCH LABORATORY

Advanced Displays and Interactive Displays Report Compendium III—Final Report

Paul Rose

ARL-SR-115

APRIL 2002

20020514 136

CAVE™ and ImmersaDesk™ are trademarks of the University of Illinois Board of Trustees.

Velcro® is a registered trademark of Velcro USA, Inc., Manchester, NH.

Windows NT® is a registered trademark of Microsoft Corporation.

Army Research Laboratory

Aberdeen Proving Ground, MD 21005-5425

ARL-SR-115   April 2002

Advanced Displays and Interactive Displays Report Compendium III—Final Report

Paul Rose, Editor
Human Research & Engineering Directorate

Abstract

This report contains more than 300 citations and abstracts of papers and presentations produced by the Advanced Displays and Interactive Displays consortium during the 5-year U.S. Army Federated Laboratory program. The program, more informally known as FedLab, was formed in 1996 to establish partnerships among the Army, industry, and academic research communities. The Advanced Displays and Interactive Displays consortium seeks to provide innovative, cost-effective solutions to information access, understanding, and management for the soldier of the future.

The research encompasses a range of topics. Some work concerns the representation of uncertainty and imprecision in databases or the representation of relationships in multimedia databases, in ways that are compatible with human cognitive-processing capabilities. Other work adopts the means of human communication (such as speech, gesture, eye gaze, and lipreading) for human-computer interaction. Additional work explores methods for incorporating information in virtual reality displays that support decision making without distracting or overwhelming the soldier. Although diverse, the research is linked by its overriding goal: the presentation of information in a form that allows effective human understanding and decision making in complex battlefield situations.

Contents

INTENTIONALLY LEFT BLANK

ADVANCED DISPLAYS AND INTERACTIVE DISPLAYS REPORT
COMPENDIUM III—FINAL REPORT

1. Introduction

The U.S. Army's 5-year Federated Laboratory (FedLab) program was created in 1996 to establish partnerships among the Army, industry, and academic research communities. Three consortia comprised the FedLab program: the Advanced Sensors consortium, the Advanced Telecommunications and Information Distribution consortium, and the Advanced Displays and Interactive Displays (ADID) consortium. Seeking to provide innovative, cost-effective solutions to information access, understanding, and management for the soldier of the future, the ADID consortium focused on cognitive related and perception-related aspects of human-computer interaction (HCI). The partners of the Displays Consortium were led by Rockwell Scientific Company[1] (RSC), an organization with wide-ranging experience in designing and developing displays for military and commercial aircraft, in integrating complex systems, and in managing complex research and development (R&D) programs. Academic institutions associated with the consortium included the University of Illinois at Urbana-Champaign (UIUC) and North Carolina Agricultural & Technical (NC A&T) State University. Much of the work at UIUC was conducted by researchers affiliated with the Beckman Institute for Advanced Science and Technology, known for its extensive program in human-computer intelligent interaction, and the National Center for Supercomputer Applications (NCSA), an institution focused on information visualization questions. Other industrial partners included Sytronics, Inc., a small business in Dayton, Ohio, which possesses a strong background in human factors research, and MCNC[2], a private, nonprofit corporation established to enhance technology-based economic development in North Carolina by providing advanced resources in electronic and information technologies to support educational and industrial institutions.

This report contains more than 300 citations and abstracts of papers and presentations produced by ADID consortium researchers during the 5-year period of FedLab's existence. The research encompasses a range of topics. Some work concerns the representation of uncertainty and imprecision in databases or the representation of relationships in multimedia databases in ways that are compatible with human cognitive-processing capabilities. Other work adopts the means of human communication (such as speech, gesture, eye gaze, and lipreading) for HCI. Additional work explores methods for incorporating information in virtual reality displays that support decision making without

---

[1] Until recently, Rockwell Scientific Company was known as Rockwell Science Center.

[2] MCNC, now known solely by its initials, was formerly known as Microelectronics Center of North Carolina.

1

distracting or overwhelming the soldier. Although diverse, the research is linked by its overriding goal: the presentation of information in a form that allows effective human understanding and decision making in complex battlefield situations.

2. Abstracts and Citations

Agre, J., Clare, L., Romanov, N., Panov, V., Kelly, J., & Klingeman, R. (2000)
Sensing positioning integrated network (SPIN): Providing situational awareness to the warfighter
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 13–17

> We describe an approach that provides situational awareness to the dismounted soldier. The fundamental challenge for our approach was to enable the warfighters to know their own location, the positions of friends, and the presence of enemies and noncombatants. This information must be conveyed via equipment that does not burden the warfighter physically or cognitively and does not depend on non-organic assets. Our solution, designated "SPIN," leverages emerging innovations in distributed microsensor networks, together with the advancing evolution of hand-held global positioning system receiver technology. The system architecture is presented, and progress is achieved through our participation in both the Advanced Displays FedLab and the Advanced Sensors FedLab.

Atchley, P., & Kramer, A. (1997)
The search for depth in the spotlight of attention
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. P), 43–52

> This experiment investigated the nature of attention in three-dimensional (3-D) space. The hypothesis of the experiment was that attention is allocated to a particular location in depth and not just to a location in x,y space. Eight observers were asked to indicate which of two target symbols appeared in a 2-by-2 matrix of boxes. The displays were stereoscopic. The boxes were placed at different locations in depth. Two of the boxes appeared near to the observer and two of the boxes appeared farther from the observer in depth. In some trials, a cue occurred at one of the four locations before the onset of the target symbol. The validity of the cue was varied. The observer's reaction time in trials where the cue indicated the incorrect x,y location but the same depth as the target was similar to trials where the cue indicated the incorrect x,y location, but the target depth was different. The findings indicate that spatial attention does not have a 3-D extent.

2

Atchley, P., & Kramer, A. F. (1998)

Spatial cuing in a stereoscopic display: Attention remains "depth-aware" with age

Journals of Gerontology: Series B: Psychological Sciences & Social Sciences, **53B**, P318–P323

Previous research has demonstrated that spatial attention is "depth aware." Reaction times (RTs) are greater for shifts in depth and two-dimensional (2-D) space than for shifts in 2-D space alone. This experiment examined whether the ability to focus attention at a depth location is maintained with advanced age. Twelve 18- to 25-year-old and twelve 62- to 85-year-old observers viewed stereoscopic displays in which one of four spatial locations was cued. Two of the locations were at a near-depth location and two were at a far-depth location. When the focus of visual attention was shifted to a new location in space (because of an invalid cue), the cost in RT for switching attention (measured as the difference between RT on valid cue and invalid cue trials) was greater when observers had to switch attention between different depth locations and different locations in 2-D space than for shifts in 2-D space alone. This effect was observed for both younger and older observers, which suggests that the ability to orient attention to a depth location is maintained with age.

Atchley, P., & Kramer, A. F. (2001)

Object- and space-based attentional selection in three-dimensional space

Visual Cognition Special Issue, 8, 1–32

It has been previously demonstrated that visual attention has an extent in depth (three-dimensional space) as well as an extent in the fronto-parallel plane (two-dimensional space). Numerous experiments have also demonstrated that attention can be allocated to objects and that "object-based" attention can overcome some of the costs associated with moving attention about in two-dimensional space. In real visual environments, objects often have an extent in depth. Four experiments were conducted to examine the nature of object-based attention in three-dimensional space. The experiments demonstrated large object-based attention benefits, as well as costs for switching attention depth. However, the costs associated with switching attention in depth were illuminated with objects that had an extent in depth. Experiments 2 through 4 examined the interaction of spatial attention in three-dimensional space and object-based attention. Evidence was found for the spread of spatial attention to objects. However, contrary to other work (Lavie & Driver, 1996), neither non-predictive exogenous spatial cues (Experiment 2) nor predictive exogenous spatial cues (Experiments 3 and 4) were able to eliminate object-based attention, which suggests that object-based attention can remain intact despite the allocation of attention spatially.

Atchley, P., Kramer, A. F., Anderson, G. J., & Theeuwes, J. (1997)
Spatial cuing in a stereoscopic display: Evidence for a "depth-aware" attentional focus
Psychonomic Bulletin and Review, 4, 524–529

Two experiments were conducted to explore whether attentional selection occurs in depth, or if the attentional focus is "depth blind," as suggested by Ghiradelli and Folk (1996). In Experiment 1, observers viewed stereoscopic displays in which one of four spatial locations was cued. Two of the locations were at a near-depth location and two were at a far-depth location; a single target was presented along with three distracters. The results indicated a larger cost in reaction time for switching attention in x,y and depth than in x,y alone, supporting a "depth-aware" attentional spotlight. In Experiment 2, no distracters were present, similar to the displays used by Ghiradelli and Folk. In this experiment, no effect for switching attention in depth was found, which indicates that the selectivity of attention in depth depends on the perceptual load imposed upon observers by the tasks and displays.

Atchley, P., Kramer, K., & Hillstrom, A. P. (2000)
Contingent capture for onsets and offsets: Attentional set for perceptual transients
Journal of Experimental Psychology: Human Perception and Performance, 26, 594–606

Four experiments were conducted to examine whether attentional set affects the ability of visual transients (onsets and offsets) to capture attention. In the experiments, visual search for an identity-defined target was conducted. In the first three experiments, the target display either onset entirely or was revealed by offsetting camouflaging line segments to reveal letters. Before the target display, there was a non-informative cue, either an onset or an offset, at one of the potential target locations. Cues that shared the same transient feature as the target display captured attention. The lack of predictable target transients led to attentional capture by all forms of transients. The final experiments with luminance changes without offsets or onsets showed attentional capture when the luminance changes were large. The results suggest that attentional set can be broadly or narrowly tuned to detect changes in luminance.

Atchley, P., Kramer, A. F., & Theeuwes, J. (1997)
Attentional control in 3-D space
Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting, 1328–1332

Two experiments investigated the nature of attention in three-dimensional space. In Experiment 1, the hypothesis that attention can be localized to a depth plane was tested. Observers searched for a red line in two arrays of green lines. The arrays of lines were near in two-dimensional space but were separated in depth. Search for the target was faster when the depth plane where the target would appear was cued,

which indicated that attention can be localized in depth. A second experiment tested the hypothesis that attending to a location in depth would reduce the effect of a distractor at other depth locations. In this experiment, search for a tilted red line was faster when a distracting vertical line was present at another depth than when it was present at the same depth as the target. Implications for display design using depth information are discussed.

Atchley, P., Kramer, A. F., & Theeuwes J. (1998)
Attentional control within three-dimensional space
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 101–106

The present study investigated whether directing attention to a particular plane in depth enables an observer to filter information from another depth plane. Observers viewed stereoscopic displays and searched for a red line segment among green line segments. Experiment 1 showed that directing attention to a particular depth plane cannot prevent attentional capture from another depth plane when the colors of the target and distractor are identical. Experiment 2 showed that directing attention to a particular depth plane can prevent attentional capture by a singleton from another depth plane when the colors of the target and distractor are different. It indicates that attentional capture by irrelevant singletons may be prevented only when both color and depth information is selective in guiding attention to the target singleton. The results suggest that retinal disparity does not have the same special status as location information in two dimensions and should be considered as just another feature along which selection may occur.

Azoz, Y., Devi, L., & Sharma, R. (1997)
Vision-based human arm tracking for gesture analysis using multimodal constraint fusion
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. P), 23–32

The use of hand gestures provides an attractive means of interacting naturally with a computer-generated display. With one or more video cameras, hand movements can potentially be interpreted by computers as meaningful gestures. One key problem in enabling such human-computer interaction without a restricted setup is the ability for the computer to localize and track the human arm in the video images. We present a technique for human arm tracking in which the arm is modeled as an articulated object that consists of rigid components. Each rigid part is assumed to give rise to a set of image features that are extracted from the video image. The tracking is performed by the assimilation of the constraints of the model and the real-time measurements incrementally to the tracking process. With this formulation, the system can handle various forms of uncertainty (e.g., image features that are missing due to

occlusion, measurement noise, etc.). Further, for reliably localizing the human hand and arm in the image, we use the multiple cues of motion, shape, and color. The image parameters for tracking the arm are then obtained by the fusing of the output of the multimodal image analysis.

Azoz, Y., Devi, L., & Sharma, R (1998)
Reliable tracking of human arm dynamics by multiple cue integration and constraint fusion
Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 905–910

> The use of hand gestures provides an attractive means of interacting naturally with a computer-generated display. In a setup using one or more video cameras, the hand movements can potentially be interpreted as meaningful gestures. One key problem in building such an interface without a restricted setup is the computer's limited ability to localize and track the human arm robustly in image sequences. This paper proposes a multiple-cue-based localization scheme combined with a tracking framework to reliably track the human arm in unconstrained environments. The localization scheme integrates the multiple cues of motion, shape, and color for locating a set of key image features. These features are tracked by a modified extended Kalman filter that uses constraint fusion and exploits the articulated structure of the arm. We also propose an interaction scheme between tracking and localization for improving the estimation process while reducing the computational requirements. The performance of the framework is validated with the help of extensive experiments and simulations.

Azoz, Y., Devi, L., & Sharma, R. (1998)
Tracking hand dynamics in unconstrained environments
Proceedings of the Third International Conference on Automatic Face and Gesture Recognition, 274–279

> A key problem in human-computer interaction via hand gestures is the computer's limited ability to localize and track the human arm in image sequences. This paper proposes a multimodal localization scheme combined with a tracking framework that exploits the articulated structure of the arm. The localization uses the multiple cues of motion, shape, and color to locate a set of image features. These features are tracked by a modified extended Kalman filter that uses constraint fusion. An interaction scheme between tracking and localization is proposed in order to improve the estimation while decreasing the computational requirement. The results of extensive simulations and experiments with real data and the contents of a large database of hand gestures involved in display control are described.

6

Baker, M. P., & Stein, R. J. (1998)
BattleView: Touring a virtual battlefield
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 54–58

> BattleView is a virtual battlefield application, developed as a research test bed for exploring the use of advanced display technologies to support user information access in large, complex, geographical information spaces. This paper describes our early work on BattleView, where we concentrated on building a flexible, responsive battle space that runs in advanced virtual environments, such as the Cave Automatic Virtual Environment (CAVE™) or ImmersaDesk™. Multimodal user interaction is supported through speech and gesture as well as point-and-click mechanisms.

Bangayan, P. T., & Chen S. L. (1999)
Noise reduction techniques for speech recognition in the military environment
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 141

> We have developed noise-reduction algorithms in an effort to improve speech recognition in noisy environments. We constructed a discrete speech recognition engine using the Entropic Hidden Markov Model Toolkit (HTK) and trained it using isolated and spelled word data from the Defense Advanced Research Projects Agency-funded remote method invocation (RMI) database. The speech samples were corrupted by additive noise obtained from personnel at the U.S. Army Research Laboratory's (ARL's) Hostile Environment Simulator (HES) and from the commercially available NOISEX database of military sounds. Results indicate that spectral subtraction reduces the error rate for stationary noise sources at signal-to-noise ratios (SNRs) ranging from 20 dB to 0 dB. However, applying spectral subtraction to non-stationary sources, which constitute many battlefield noises, resulted in an increased error rate. To mitigate the problem of non-stationary noise, a dual microphone approach has been taken. We filtered a signal consisting of both speech and noise by using a second signal consisting of noise alone. Data were collected at Rockwell Scientific Company; the noise mix was provided by ARL HES personnel. The data collection constituted a first step toward an audio-visual database for bimodal speech recognition planned for Fiscal Year 1999. Samples of the audio-only data collection are presented.

Banks, R., Wickens, C. D., & Hah, S. (1998)
Commander's display of terrain information: Manipulations of display dimensionality and frame of reference to support battlefield visualization

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 69–73

> We assessed the effects of three battlefield perspectives on terrain visualization. Thirty Army officers answered a series of battlefield questions while viewing the battlefield on electronic displays that gave three perspectives: A two-dimensional vertical perspective giving the subjects a contoured map view; a three-dimensional elevated perspective presenting the same terrain from a 45-degree viewing angle, which could be rendered at the subject's choice with either contour lines or shadowing; and a three-dimensional interactive immersive perspective, which allowed subjects to select a location on the surface of the battlefield, travel to it, and rotate their viewpoint of the battlefield. Results indicated that distance questions were best answered with the two-dimensional map view and that line-of-sight visibility questions were most accurately supported by the immersive perspective, although with a time cost. Questions concerning troop mobility were supported equally by all three viewpoints. Subjects' performance was correlated with spatial abilities, and those subjects with lower spatial abilities compensated by using the interactive immersive display more frequently. The findings support the importance of multiple viewpoints and task analysis in battlefield display design.

Bargar, R. (1997)
Generating and controlling synchronous sound for interactive graphical computing environments

Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 53–62

> This paper introduces concepts, implementation challenges, and solution strategies for generating synchronous sound with computer graphics in real time. Benchmarks and constraints of high-fidelity audio signal processing are introduced. Diagnoses are provided of existing hardware and software subsystems. An architecture is presented for achieving synchronous real-time synthesis of high-fidelity audio signals.

Bargar, R., & Choi, I. (1998)
Sonification of probabilistic belief networks

Proceedings of the 1998 IEEE International Conference on Systems, Man, and Cybernetics, 1, 1020–1025

> We describe a working sonification system for design and implementation of real-time data-driven auditory displays. Sonification is applied to enhance the visual display of an interactive decision support system consulting a Bayesian belief network. The sonification case presented in this paper employs the concept of an auditory signature. The

auditory signature is attributed to the nodes that observers wish to keep track of, particularly for monitoring the dynamics of internal nodes. The objective is to provide fine gradients of auditory information to help observers be aware of the relative contribution of internal nodes to the final outcome. For implementation of the prototype system, we developed a task-based model of Bayesian belief network dynamics. This model provides criteria for the design of a sonification architecture. The early development of a prototype architecture allows the research team to identify constraints presented by the visual display and interactivity of the Bayesian belief network and to develop alternatives early in the project cycle.

Bargar, R., & Choi, I. (1999)
Sonification of dynamic data representation networks to reduce visual overload and enhance situational awareness
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 21–25

> We describe a working sonification system for design and implementation of real-time data-driven auditory display. Sonification is applied to enhance the visual display of an interactive decision support system consulting a Bayesian belief network. The sonification case presented in this paper employs the concept of auditory signature. The auditory signature is attributed to the nodes that observers wish to keep track of, particularly for monitoring the dynamics of internal nodes. The objective is to provide fine gradients of auditory information to help observers be aware of the relative contribution of internal nodes to the outcome. For implementation of the prototype system, we developed a task-based model of Bayesian belief network dynamics. This model provides criteria for the design of sonification architecture. The early development of prototype architecture allows the research team to identify constraints presented by the visual display and interactivity of the Bayesian belief network and to develop alternatives early in the project cycle.

Bargar, R., Choi, I., & Betts, A. (1999)
Scoregraph: A software architecture for rapid configuration of multimodal interaction in distributed virtual environments
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 41–45

> This paper presents software architecture for rapid configuration of multidimensional and multimodal interactions in virtual environments. The architecture is currently in active use in the Integrated Support Laboratory, Beckman Institute, UIUC. Observation is described as interaction with a virtual environment to extract information in a time-critical manner. In the present research, an observer's multimodal capacity is supported by time scheduling techniques for parallel

processing of sensors and displays to provide synchronous perceptual feedback. This modality is coupled to multidimensional numerical simulations. The ScoreGraph software architecture facilitates a temporal framework for dynamic interplay in virtual environments. A temporal framework is complementary to the static spatial organization of geometric graphical objects. Design criteria include the management of computing resources, a configuration of an observation space, and virtual reality (VR) authoring. The temporal criteria in VR authoring have to do with efficient reconfiguration of interactive capacity in a virtual scene and the dynamics of services exchanged among parallel processes.

Barnes, M. J., & Fichtl, T. (1999)
Cognitive issues for the intelligence analyst of the future
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 15–19

> The purpose of the paper is threefold: (1) identify important trends that affect the future analyst, (2) discuss the cognitive implications of these trends, and (3) suggest empirical and theoretical issues for further research. Four important cognitive areas are discussed in detail: knowledge acquisition, situation awareness, prediction, and intuitive processes. The conclusion is that the 21st century analyst will face radically new technology and a variety of unconventional intelligence missions. Research and decision support are discussed as possible amelioratives.

Barnes, M. J., & Knapp, B. G. (1997)
Collaborative planning aids for dispersed decision making at the brigade level
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 63–72

> The purpose of the project is to understand the effects of dispersion on the brigade planning process. Cognitive theory related to information presentation and knowledge representation was discussed. Optimal presentation was found to depend on the combat role being performed. The implication was that command and control involved diverse combat views that were distinct but connected. The problems that this diversity would have for dispersed operations and a possible experimental framework were discussed as well. ARL modeling efforts and the supporting human performance experimentation paradigms were identified to isolate causal factors for hypothesized performance decrements. The same methods were also suggested as a means to develop and to evaluate solutions for identified problems.

Barnes, M. J., Sohn, Y. W., & Doane, S. (2000)
Modeling the intuitive warfighter
*Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium,* 163

> This research focuses on the cognitive processes of military planners in
> novel combat situations by investigating the extent to which ADAPT, a
> computational model of human planning, can be used to provide insight
> into the cognitive operations that take place during human battlefield
> planning tasks. The long-range goal of this program is to construct
> quantitative models of battlefield planner's knowledge that predict
> warfighter ability to detect deception and react to unforeseen changes in a
> battle-planning-task context as a function of expertise. We refer to expert
> personnel with this ability as "intuitive" warfighters. Our
> accomplishments include modeling warfighter problem solving in both
> conventional and non-conventional combat scenarios using a hybrid
> symbolic-connectionist architecture.

Barnes, M. J., & Wickens, C. D. (1998)
*The commander's ability to visualize battlespaces: A multi-view
approach*
*Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium,* 1–5

> Modern battle spaces are complex, information intensive, and extremely
> fast paced. The problem this paper addresses is the human element of the
> battle space visualization process. Visualization refers to both graphic
> representations and mental images of complex processes (Barnes, 1997;
> Wickens, Merwin, & Lin, 1994). The purpose of battle space visualization
> is to enhance the commander's (or his staff's) ability to understand the
> unfolding battle in order to make timely and informed tactical decisions.
> This goal not only subsumes terrain visualization and viewing troop
> deployments but also implies an intuitive understanding of the battle
> process, including the visualization of possible end states and their
> consequences (Barnes, 1995; Beseler, 1997). Our focus is the behavioral
> link between different representation techniques and the human's ability
> to better understand and make decisions about the battle process. We
> discuss some theoretical notions related to decision making and to
> perception; then we discuss recent empirical results and their import for
> understanding human visualization. Finally, we suggest future areas of
> investigation that we feel will lead to principles that will support a multi-
> view visualization system.

Barnes, M. J., Wickens, C. D., & Smith M. (2000)
Visualizing uncertainty in an automated national missile defense (NMD) simulation environment

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 107–111

> Two experiments were conducted to investigate the effects of various information presentation variables on risk perception in a highly automated missile defense simulation. The results of the first experiment indicated superior situation awareness when risk was presented as the expected frequency of leaker missiles rather than as abstract probabilities. However, the way of presenting risk had no effect on decisions to remove interceptors from reserve status. There was, however, an important interaction between immediate and delayed risk. In contrast, the second experiment showed a significant effect on reserve decisions, depending on when the negative battle events occurred (both primacy and recency effects), but showed no effect when various risk measures were highlighted. The results were discussed in terms of cognitive biases and their implications for future display designs.

Beebe, D., & Tang, H. (1997)
A tactile chording system for the dismounted soldier

Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. P), 33–41

> A rifle-mounted tactile interface is proposed and the chording system is implemented for demonstration and for conducting human experiments. The interface is intended for use by the dismounted soldier. A typical scenario is discussed. The principles of operation of the chording system, which use conductive polymer sensors as multi-state input elements, are described, including basic multi-state concepts, pressure sensor operation, scaling methods, and feedback. The system design and current implementation are presented.

Behringer, R. (1998, October)
Improving the precision of registration for augmented reality in an outdoor scenario by visual horizon silhouette matching

paper presented at the *International Workshop on Augmented Reality,* San Francisco, CA

> A system for enhancing situation awareness in an outdoor scenario is being developed. The goal of such a system is to provide information through an overlay superimposed onto a video stream or directly into a head-mounted display; the superposition is done by augmented reality techniques. In an outdoor scenario, the registration between the overlay and the real world can be obtained by a combination of global positioning system, digital compass, and inertial sensors. However, these methods lack the precision that is required for a convincing augmented reality overlay. A means to increase the registration precision, if the terrain is well structured, is to exploit the known position of visual terrain features

or man-made objects. If visible, the horizon silhouette provides cues for observer orientation. In a first step toward a system for visual outdoor registration, visual registration through horizon silhouettes has been demonstrated on single-image snapshots. The theoretical 360° horizon silhouette could be computed from U.S. Geological Survey digital elevation maps, which provide a grid of elevation data. The best match of the extracted visible silhouette segment onto the predicted 360° silhouette provides orientation (elevation, azimuth) and calibration of the observer camera. The system runs on a PC (200 MHz) and is being ported to a wearable platform.

Behringer, R. (1999)
A hybrid registration system for outdoor augmented reality
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 117–120

> Using augmented reality (AR) to enhance the soldier's situation awareness requires registration of the displayed information with the real world. A hybrid registration system has been developed for registration in an outdoor environment. The system consists of the following components: a magnetometer for determining magnetic north (digital compass) and an inclinometer for obtaining the user's head tilt and roll angle. An additional visual silhouette registration system using a camera, aligned with the user's view, improves the accuracy of the orientation. The system is prepared for later integration with a global positioning system receiver for obtaining location. The registration system is being ported to a van, which will allow it to be tested at arbitrary locations. It is also being ported to a mobile wearable PC, which can provide simple AR functionality. The AR system will be capable of providing remote AR to a central command post. The paper describes the system architecture and presents first results of the overlay.

Behringer, R. (1999)
Improved registration precision through visual horizon silhouette matching
in Behringer, R., Klinker, G., & Mizell, D. W. (Eds.). *Augmented reality: Placing artificial objects in real scenes* (pp 225–232). Natick, MA: A. K. Peters

> The registration precision of an augmented reality system for enhancing the situation awareness in an outdoor setting can be improved by the use of visual clues. Terrain silhouettes can provide unique features to be matched with digital elevation map (DEM) data. The best match of a visually extracted silhouette with the DEM silhouette provides camera/observer orientation (elevation and azimuth angle). We have developed such a registration system, which runs on a PC (Pentium Pro, 200 MHz) and is being ported to a wearable AR system.

Behringer, R. (1999)
A system for inertial stabilization of a video display
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 127–131

> In the future, soldiers will operate equipment while riding over rough terrain in the U.S. Army's command and control vehicle (C2V). The motion-induced vibration in this environment causes a massive reduction in the readability of the displays. To mitigate this problem, we have developed a system that can compensate for computer monitor motion by projecting the information onto the display in an inertially stabilized window. The window is shifted on the monitor in the opposite direction as the monitor motion. A three-axis linear accelerometer measures the acceleration at the display. The acceleration data are used to shift the display window so that it appears at a fixed spatial location, although the monitor itself is moving. The system is implemented on a standard PC (Pentium Pro, 200 MHz, Windows NT® 4.0) using commercial off-the-shelf components. In the paper, we present an overview of the algorithm, the system implementation, and results from vibration simulation.

Behringer, R. (2001)
Stabilization of a display in a moving environment
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 131–134

> In a highly dynamic environment such as a moving vehicle, display readability is reduced by the vibrations caused by the ride over rough terrain. Since the display monitors are mechanically connected with the vehicle, they perform the same vibration motions as the vehicle. Stabilization of the display monitor through mechanical dampening devices is expensive, and retrofitting existing installations often is not possible. Funded by the FedLab Consortium, Rockwell Scientific Company has developed a prototype of a system that improves display readability by stabilizing the display content instead of stabilizing the monitor. This is done by software, without the need for an expensive hardware installation. An accelerometer sensor attached near the monitor captures the motion of the monitor. Integrating these data provides a measure for the absolute displacement in inertial space. This displacement is used to shift the display content opposite to the monitor movement to achieve an inertially stable visualization. In order to prevent the display's content moving from "out of the display area," a relaxation mechanism has been implemented to pull the display content back to the center of the monitor in the absence of acceleration. This approach is well suited for oscillating motions in the range of 2 Hz to 10 Hz. The lower boundary is determined by the magnitude of the displacement and the precision of the acceleration data; errors in the acceleration have a large influence on the calculated displacement because of double integration. The upper boundary is determined by system latency and measurement

sampling rate. Such a "virtual inertial display stabilization system" can keep the display contents more stable relative to the user. Scenarios of such a system are applied whenever users of displays in a highly dynamic environment are subjected to vibrations.

Behringer, R., & Ahuja, N. (1998)
Image registration for computer vision-based augmented reality
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 44–48

Displays that provide graphical information as an overlay on the view of the real world can substantially increase situation awareness by augmenting the visible world with additional relevant information. Such displays can be either see-through head-mounted displays carried by the soldier on the battlefield or conventional monitors in a command and control center showing a video image of the real world, which is augmented by additional graphical information. The information to be displayed in such an augmented display must be superimposed with objects in the real world. Because of the three-dimensional nature of this problem, a simple two-dimensional image registration approach (as used in a satellite image registration system) is not suitable. The problem that has to be addressed in order to correctly interpret the spatial coherence is the estimation of the viewer's position, orientation, and motion. Important visual cues for obtaining these parameters are the surface silhouettes of real-world objects. In a well-structured terrain, the horizon silhouette formed by the terrain shape is a distinct feature that characterizes the viewpoint. The unique shape of this silhouette can be exploited to obtain the orientation of the camera at a known position. We have developed such a system for image and world registration, based on matching an extracted video silhouette segment with a pre-computed 360° silhouette profile from a digital elevation map. Results are shown for registration of the mountainous regions around Thousand Oaks, California.

Behringer, R., Klinker, G., & Mizell, D. W. (Eds.). (1999)
Augmented reality: Placing artificial objects in real scenes
Natick, MA: A. K. Peters

This book contains papers presented at the International Workshop on Augmented Reality. Augmented reality (AR) typically consists of computer-generated information displayed on a transparent helmet-mounted display and superimposed on real-world surroundings, but the information can assume other forms such as sound. Applications discussed in the book include industrial manufacturing in an airplane factory; virtual prototyping, in which a product can be envisioned in its real-world surroundings and can be tested for design flaws; error diagnostics and maintenance of complex machinery, in which status information and instructions can be superimposed on the critical machinery component; and enhancement of situation awareness and

perception of the real world, achieved by placing virtual objects or information cues in the real world. In a sense, AR occupies an intermediate position between virtual reality and the real world. Academic and industrial researchers have different views of AR. Academics aim at applying AR to daily life in a user-centric paradigm, often implementing AR in a kind of wearable system. Industrial researchers are more interested in adopting current technology for task-specific industrial processes and are less interested in providing the most intuitive and advanced interaction with the application.

Behringer, R., Seagull, J., & Wickens, C. (1998)
Human perception under vibration in the C2V

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 115–119

Vibrations during vehicle ride degrade the performance of human operators in the vehicle. In command and control vehicles (C2Vs), the legibility of a display is reduced significantly by the vibration. In order to develop measures to compensate for this degrading vibration effect, we measured the power spectral density of the vibration of a moving vehicle. The source of low-frequency vibration is the roughness of the terrain on which a vehicle is moving. The vehicle movement is shaped by the dynamic system of vehicle suspension. In conventional four-wheeled vehicles, two distance peaks near 1 Hz and 10 Hz determine the vibration characteristics. Measurements in an Army C2V indicate that these major resonance peaks are at frequencies beyond 20 Hz. Below 20 Hz, however, there is a significant power spectral density that contributes to the observed degradation of human operator perception. A preliminary analysis of the vertical, lateral, and transverse power spectral density functions of the ride vibrations in the C2V is presented, and the impact of these vibrations on the human operator's perception is discussed.

Behringer, R., Tam, C., McGee, J., Sundareswaran, S., & Vassiliou, M. (2000)
A wearable augmented reality testbed for navigation and control

Proceedings of IEEE and ACM International Symposium on Augmented Reality, 12–19

Personal applications employing augmented reality (AR) for information systems require ease of use and wearablility. Progress in hardware miniaturization is enabling the development of wearable test beds for such applications and is providing sufficient computing power for the demanding AR tasks. Rockwell Scientific Company has assembled a wearable test bed for AR applications, comprised entirely of commercial off-the-shelf hardware components. The system is designed to be worn like a jacket, with all hardware attached and affixed to a vest frame (Xybernaut) with concealed routing of cables under Velcro© channels. Two possible configurations allow the system to be used either in a stand-alone mode (itWARNS) or to be linked to a larger scale multi-modal user interface test bed (WIMMIS). Completely tetherless operation is made possible by wireless digital connections as well as analog video and three-

dimensional audio connections over radio frequencies. This paper describes these two test bed configurations as well as some of the AR applications developed on this test bed.

Berry, G. A., Pavlovic, V., & Huang, T. S. (1998, November)
BattleView: A multimodal HCII research application

paper presented at the *Proceedings of the Workshop on Perceptual User Interfaces*, San Francisco, CA

> To demonstrate some of our research topics in human-computer intelligent interaction (HCII), we employ two modes of natural human-computer interaction to control a virtual environment. By using speech and gesture recognition, we outline the control of a virtual environment research test bed (BattleView) without the need for traditional virtual reality interfaces such as a wand, mouse, or keyboard. The use of features from both speech and gesture creates a unique interface where different modalities complement each other in a more "human" communication style.

Cantú-Paz, E. (1999, July)
Migration policies, selection pressure, and parallel evolutionary algorithms

paper presented at the *Late-Breaking Papers of the 1999 Genetic and Evolutionary Computation Conference*, Orlando, FL

> This paper investigates how the policy used to select migrants and replacements affects the selection pressure in parallel evolutionary algorithms (EAs) with multiple populations. The four possible combinations of random and fitness-based emigration and replacement of existing individuals are considered. The investigation follows two approaches. The first is to calculate the "take-over" time under the four migration policies. This approach makes several simplifying assumptions, but the qualitative conclusions that are derived from the calculations are confirmed by the second approach. The second approach consists of quantifying the increase in the selection intensity. The results may help to avoid excessively high (or low) selection pressures that may cause the search to fail and may offer a plausible explanation of the frequent claims of super linear increases in the execution rate of parallel EAs.

Cantú-Paz, E. (1999)
Topologies, migration rates, and multi-population parallel genetic algorithms

Proceedings of Genetic Algorithms and Classifier Systems, 91–98

> This paper presents a study of parallel genetic algorithms (GAs) with multiple populations (also called *demes* or *islands*). The study makes explicit the relation between the probability of reaching a desired solution with the deme size, the migration rate, and the degree of the connectivity graph. The paper considers arbitrary topologies with a fixed number of

neighbors per deme. The demes evolve in isolation until each converges to a unique solution. Then the demes exchange an arbitrary number of individuals and restart their execution. An accurate deme-sizing equation is derived, and it is used to determine the optimal configuration of an arbitrary number of demes that minimizes the execution time of the parallel GA.

Cepeda, N. J., & Kramer, A. F. (1999)
Strategic effects on object-based attentional selection
Acta Psychologica, **103**, 1–19

The same-object benefit (i.e., faster and/or more accurate performance when two target properties to be identified appear on one object than when each of the properties appears on different objects) has been a robust and theoretically important finding in the study of attentional selection. Indeed, the same-object benefit has been interpreted to suggest that attention can be used to select objects and perceptual groups rather than unparsed regions of visual space. This article reports and explores a different-object benefit (i.e., faster identification performance when two target properties appear on different objects than when they appear on a single object). Participants in all three experiments included 7 male and 37 female 18- to 31-year-old college students. The results from the three experiments suggest that the different-object benefit was the result of mental rotation and translation strategies that participants performed on objects in an effort to determine whether two target properties matched or mismatched. These image manipulation strategies appear to be performed with similar but not with dissimilar target properties. The results are discussed in terms of their implications for the study of object-based attentional selection.

Chakrabarti, K., & Mehrotra, S. (1997)
Concurrency control in R-trees
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. P), 1–14

The nature and types of information in a dynamic battlefield environment include geo-referenced satellite images and terrain elevation; maps containing terrain features such as roads, enemy unit deployment, activities and targets; and spatio-temoporal objects such as logistic, tactical, and collection management plans. Efficient processing of queries about such objects in a database requires support of access paths using an effective multidimensional data structure. While a large body of research exists about multidimensional data structures (grid files, R-trees, hB-trees, to name a few), none of the data structures themselves have been fully integrated into any commercial strength database management system, partly because of the lack of effective techniques to support concurrent access to and manipulations of these data structures. This paper identifies problems in supporting concurrent operations over multidimensional data structures and sketches solutions in the context of R-trees. The R-tree

is a popular multidimensional data structure that has been incorporated into Illustra's data management system and has also been implemented at part of the Shore object management system.

Chakrabarti, K., & Mehrotra, S. (1998)
Dynamic granular locking approach to phantom protection in R-trees
Proceedings of the 14th International Conference on Data Engineering, 446–454

> Over the last decade, the R-tree has emerged as one of the most robust access methods for multidimensional databases. However, before the R-tree can be integrated as an access method in a commercial strength data management system, efficient techniques for transactional access to data via R-trees need to be developed. Concurrent access to data through a multidimensional data structure introduces the problem of protecting ranges specified in the retrieval from phantom insertion and deletions (phantom problem). Existing approaches to phantom protection in B-trees (namely, key-range locking) cannot be applied to multidimensional data structures since they rely on a total order over the key space on which the B-tree is designed. This paper presents a dynamic granular locking approach to phantom protection in R-trees. To the best of our knowledge, this paper provides the first solution to the phantom problem in multidimensional access methods based on granular locking.

Chakrabarti, K., & Mehrotra, S. (1999)
Efficient concurrency control in multidimensional access methods
SIGMOD Record, 28(2), 25–36

> The importance of multidimensional index structures to numerous emerging database applications is well established. However, before these index structures can be supported as access methods in a commercial strength database management system (DBMS), efficient techniques to provide transactional access to data via the index structure must be developed. Concurrent access to data via index structures introduces the problem of protecting ranges specified in the retrieval from phantom insertions and deletions (the *phantom problem*). This paper presents a dynamic granular locking approach to phantom protection in Generalized Search Trees (GiSTs), an index structure supporting an extensible set of queries and data types. GiSTs provide a set of interfaces using a new multidimensional index structure that can easily be integrated into a DBMS. The granular locking technique offers a high degree of concurrency and has a low lock overhead. Through our experiments, we show that the technique scales well under various system loads. Since a wide variety of multidimensional index structures can be implemented with GiST, the developed algorithms provide a general solution to concurrency control in multidimensional access methods. To the best of our knowledge, this paper provides the first such solution based on granular locking.

Chakrabarti, K., & Mehrotra, S. (2000, September)
Local dimensionality reduction: A new approach to indexing high dimensional spaces

Paper presented at the 26th International Conference on Very Large Databases, Cairo, Egypt

Many emerging application domains require database systems to support efficient access over highly multidimensional data sets. The current state-of-the-art technique for indexing high-dimensional data is to first reduce the dimensionality of the data via principal components analysis and then index the reduced dimensionality space via a multidimensional index structure. This technique, referred to as global dimensionality reduction (GDR), works well when the data set is globally correlated (i.e., when most of the variation in the data can be captured by a few dimensions). In practice, however, data sets are often not globally correlated. In such cases, reducing the data dimensionality via GDR causes significant loss of distance information, resulting in a large number of false positives and, thus, a high query cost. Even when a global correlation does not exist, subsets of data that are locally correlated may exist. In this paper, we propose a technique called local dimensionality reduction (LDR) that tries to find local correlations in the data and performs dimensionality reduction on the locally correlated clusters of data individually. We develop an index structure that exploits the correlated clusters to efficiently support point, range, and k-nearest neighbor queries over high-dimensional data sets. Our experiments on synthetic as well as real-life data sets show that our technique (1) reduces the dimensionality of the data with significantly lower loss in distance information compared to GDR and (2) significantly outperforms the GDR, original space indexing, and linear scan techniques, in terms of the query cost for both synthetic and real-life data sets.

Chakrabarti, K., Mehrotra, S., Ortega, M., Porkaew, K., & Winkler, R. (1998)
Processing uncertainty queries in database management systems

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 20–26

Emerging applications (including many military applications) require explicit mechanisms to represent and process uncertainty in queries and in the data stored in databases. Most current approaches for supporting uncertainty inquires layer a reasoning component on top of existing relational database management systems (DBMSs), which resolves the uncertainty in queries outside the DBMS. While the layered approach is attractive because of its simplicity, and since it requires minimal extensions of existing DBMS technology, it has some fundamental shortcomings that limit its usefulness to simplistic applications. This paper proposes an extended relational model together with a suitably extended relational algebra as an alternative mechanism to incorporating uncertainty in queries. In contrast to the layered approach, the proposed

model allows uncertainty to permeate database processing, overcoming many of its limitations. The paper identifies challenging research issues that we are currently addressing in developing the proposed framework.

Chakrabarti, K., Ortega-Binderberger, M., Porkaew, K., & Mehrotra, S. (2000)
Similar shape retrieval in MARS

Proceedings of the IEEE International Conference on Multimedia and Expo, 2, 709–712

This paper presents a novel approach for representing two-dimensional shapes, which adaptively models different portions of the shape at different resolutions that have higher resolutions where it improves the quality of the representation and lower resolution elsewhere. The proposed representation is invariant to scale, translation, and rotation. The representation is amenable to indexing via existing multidimensional index structures and can thus support efficient similarity retrieval. Our experiments show that the adaptive resolution technique performs significantly better, compared to the fixed resolution approach previously proposed in the literature.

Chakrabarti, K., Porkaew, K., & Mehrotra, S. (2000)
Efficient query refinement in multimedia databases

Proceedings of the 16th International Conference on Data Engineering, 196

We describe a method of searching database management systems (DBMS), based on query refinement, that is, a search technique that allows users to interactively specify their informational needs to the system by providing relevance ranking on examples of objects. Rather than treat each refined query as a "starting" query, alternate approaches are explored that significantly improve the cost of evaluating refined queries by exploiting the observation that the refined queries are not modified drastically from one iteration to another. As a result, most of the execution cost can be saved by appropriately exploiting the information generated during the previous iterations of the query. The technique is applicable to DBMS containing multimedia objects (e.g., images, video, audio, time series, spatial, and spatio-temporal data). Our experiments over a large image/text collection (COREL dataset) show that the proposed techniques provide significant improvements in performance.

Chakrabarti, K., Porkaew, K., & Mehrotra, S. (2000, September)
Refining top-k selection queries based on user feedback

Paper presented at the 26th International Conference on Very Large Databases, Cairo, Egypt

In many applications, users specify target values for certain attributes or features, without requiring exact matches to these values in return. Instead, the result is typically a ranked list of "top k" objects that best matches the specified feature values. User subjectivity is an important aspect of such queries; that is, which objects are relevant to the user and which are not depends on the perception of the user. Because of the subjective nature of top-k queries, the answers returned by the system to

a user query often do not immediately satisfy the user's need for a variety of reasons such as the weights and the distance functions associated with the features do not accurately capture the user's perception, or the specified target values do not fully capture his or her informational need. In such cases, the user would like to refine the query and resubmit it in order to receive a better set of answers. While much research has been conducted on query refinement models, we are not aware of any work supporting refinement of top-k queries efficiently in a database system. Done naïvely, each refined query can be treated as a starting query and can be evaluated from scratch. This paper explores alternate approaches that significantly reduce the cost of evaluating refined queries by exploiting the observation that the refined queries are not modified drastically from one iteration to another. Our experiments over a real-life multimedia data set show that the proposed techniques save more than 80% of the execution cost of refined queries over the naïve approach and are more than an order of magnitude faster than a simple sequence scan.

Chan, M. T. (1999)
Tracking lip motion at video rate for bimodal speech recognition
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 47–50

In support of the development of a vision-assisted speech recognition system, we have developed a video-based algorithm that can track movements of the speaker's lips during speech utterances. The method takes advantage of prior knowledge that we have about the shape of the speaker's lips and their color in contrast to that of the skin. Because the method (a) uses an explicit coarse-to-fine local search strategy, (b) constrains deformation of the model from its reference shape in an affine subspace, and (c) monitors errors and ignores outlier measurements as necessary, the algorithm is robust but still runs at a real-time video rate. Using a fast lip localization algorithm based on clustering analysis that uses the hue and saturation images, our system can also self-start without requiring user intervention at run time. We plan to incorporate the tracking subsystem into a real-time bimodal speech recognition system.

Chan, M. T. (1999)
Visual speech interface: Apparatus and algorithms
1999 World Aviation Congress, Society of Automotive Engineers (Report No. 99WAC-150)

To make speech recognition a viable input modality in the cockpit, we propose to include visual speech input to improve robustness of the approach in the presence of noise. The visual speech interface includes a head-mounted lip imaging apparatus and algorithms to recognize spoken words visually. Our algorithms are based on a few components that address all issues related to lip localization, lip shape model extraction, tracking, and feature extraction and recognition. We demonstrate the

practicability of the concept with a visual speech recognizer for a discrete word recognition task that is relatively simple but achievable in real time.

Chan, M. T. (2001)
A hybrid visual processing front end for improved speech recognition
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 91–94

> A good front end for visual feature extraction is an important element of any audio-visual speech recognition system. We propose a new visual feature representation that combines feature extraction from two sources: the geometric features of the lips and the flesh tones around the lips. Using a contour-based, lip-tracking algorithm, geometric features, including the height and width of the lips, are extracted. Rather than attempt to extract all the information from the area above and below the lips, a subset of all possible pixels was selected to minimize computational requirements without losing significant amounts of detail. The pixels were selected in reference to the tracked boundary of the upper and lower lips. Boundary tracking allows for proper scaling of the pixel-based feature vector to one of constant length. We show the advantage of the combination of these features for visual speech recognition.

Chan, M. T., Zhang, Y., & Huang, T. S. (1998)
Integrating visual and acoustic features for speech recognition
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 64–68

> We investigate how visual features derived from the speaker's lip movements can augment acoustic speech signals to improve the accuracy of an automatic speech recognition system. Our current test bed includes a system that tracks in real time positions of color markers placed on the speaker's lips while utterances are simultaneously recorded. Normalized vertical and horizontal openings of the mouth are used to augment the standard acoustic features to train our continuous speech recognizer, which is based on HMMs. We employ context-dependent models to capture co-articulation effects present in both the acoustic and the visual measurements. Two different schemes for fusing information from the two different sources were investigated. We found that a bimodal recognizer outperforms an acoustic-only recognizer in the presence of acoustic noise and does even better at low SNRs.

Chan, M.T., Zhang, Y., & Huang, T. S. (1998)
Real-time lip tracking and bimodal continuous speech recognition
Proceedings of the 1998 IEEE Second Workshop on Multimedia Signal Processing, 65–70

> We investigate a bimodal approach to improve the accuracy of an automatic speech recognition system by augmenting acoustic speech features with visual features derived from the lip movement of the

speaker. Our initial test bed includes a system that tracks in real time the positions of color markers placed on the speaker's lips while utterances are simultaneously recorded. By combining both features, we "train" a context-dependent hidden Markov model-based recognizer using continuous speech data that we collected, based on a confined vocabulary useful for our application area. Our preliminary results show that the experimental bimodal recognizer has a higher recognition accuracy than the acoustic-only counterpart, especially at low signal-to-noise ratios. We are currently incorporating into our recognizer a new algorithm for lip tracking so that markers would not be needed. Currently, the algorithm can track the outline of the lips in real time with some moderate assumptions about the speaker.

Chen, S. L. (1998)
Improving the accuracy of speaker-independent hidden Markov model-based speech recognition with redundant segregated models
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 85–89

This paper discusses the results of an experiment that demonstrates improved accuracy in automatic speech recognition via a hidden Markov model (HMM). By relying on multiple, redundant, demographically and environmentally segregated speech models, the speech recognizer is speaker independent. Traditional approaches to speaker-independent automatic speech recognition using HMMs rely on derived statistics of speech feature vectors from a large aggregated set of training speakers. Members of the training set are chosen so that their demographic composition (including dialects, genders, and ages) closely aligns with the demographics of the target set of end users. Because of the diversity of human speech characteristics, the full aggregation of a demographically representative set of speakers may produce excessive variance in statistical speech models, leading to possible overlap of models and substitution errors when speakers attempt to use the trained speech recognizer. A similar argument can be made regarding aggregation of speech data collected in widely varying acoustic environments. Our experiments investigate the merit of training multiple redundant models of selected U.S. English sub-words, in which each of the constituent member-models in a redundant set has been trained on an exclusive subset of the overall training data. These mutually exclusive subsets are partitioned by demographic or environmental criteria and together comprise the overall training population. Using these more narrowly focused training sets can potentially produce tighter variances in the resulting models and can improve speech recognition accuracy, especially when information concerning the demographic characteristics of a given speaker or the ambient acoustic environment are specified *a priori*. When speaker or acoustic information is given *a priori*, the described techniques are shown to improve recognition rates by as much

as 10%, but even if no information is given *a priori*, recognition rates improve by as much as 2.5%.

Chen, S., & Chan, M. (1997)
Challenges in robust automatic speech recognition
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 43–51

This paper discusses accuracy problems often encountered in current commercial speech recognition technologies and proposes several modifications of traditional speech recognition approaches, which might improve recognizer performance. Automatic speech recognition is a desirable component of many user interfaces because of its inherent flexibility and low learning curve, but the technology often suffers from low-accuracy problems, which make its behavior more probabilistic than deterministic, thus negating its usefulness as a control method. Among the challenging factors we must consider are changes in the user's speech induced by perceived stress, overload of microphones because of excessive loudness on the battlefield, dynamic and powerful ambient noise leading to low or even negative signal-to-noise ratios, noise characteristics that mask traditional acoustic features used to classify speech, and statistical speech models that do not accurately reflect the user's characteristics or the environment's. Potential approaches to alleviating these problems include the use of microphones with greater dynamic range and durability; the use of two or more microphones to help distinguish user speech from ambient noise, active noise cancellation or reduction; the use of acoustic feature sets that are resistant to noise and stress, explicit modeling of known battlefield sounds, building statistical models of human speech in noisy and stressful environments; and the use of machine vision techniques to recover articulatory features of speech that may be difficult or impossible to detect by acoustic means in battlefield environments.

Chernyshenko, O., & Sniezek, J. A. (1998, November)
Priming for expertise and confidence in choice: Evaluating the global improves calibration for the specific
paper presented at the annual meeting of the *Judgment and Decision-Making Society*, Dallas, TX

Two experiments investigated the relationship between expertise priming and subjects' over- or under-confidence in their judgments. Judgment about an event is based on an individual's subjective estimate of an event's probability of occurrence. During high uncertainty, for example, subjective probabilities often exceed the actual probability of an event, leading to over-confidence in one's judgment. Over-confidence was reduced when decisions were difficult, and under-confidence was reduced when they were easy if subjects were guided through an exercise

that focused attention on their beliefs about their expertise (i.e., when the subjects were "primed " for expertise).

Chu, S. M., & Huang, T. S. (2001)
Improving bimodal speech recognition using coupled hidden Markov models
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 85–89

> In this paper, we present a bimodal speech recognition system in which the audio and visual modalities are modeled and integrated with coupled hidden Markov models (CHMMs). CHMMs are probabilistic inference graphs that have hidden Markov models as sub-graphs. Chains in the corresponding inference graph are coupled through matrices of conditional probabilities modeling temporal influences between their hidden state variables. The coupling probabilities are both cross chain and cross time. The latter is essential for allowing temporal influences between chains, which is important in modeling bimodal speech. Our bimodal speech recognition system employs a two-chain CHMM, with one chain associated with the acoustic features and the other with the visual features. A deterministic approximation for maximum *a posteriori* (MAP) estimation is used to enable fast classification and parameter estimation. We evaluated the system on a speaker-independent connected digit task. Compared with an acoustic-only automatic speech recognition system "trained" with only the audio channel of the same database, the bimodal system consistently demonstrates improved noise robustness at all signal-to-noise ratios. We further compare the CHMM system reported in this paper with our earlier bimodal speech recognition system in which the two modalities are fused by concatenating the audio and visual features. The recognition results clearly show the advantages of the CHMM framework in the context of bimodal speech recognition.

Cibulskis, M. J. & DeJong, G. (1999)
Interfaces that learn: Path planning through minefields
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 143

> An approach is described for studying the problems involved in implementing an adaptable human-computer interface. To provide useful information and guidance, an adaptable interface must be sensitive to the expertise level of the user and to the user's tolerance to system interference, which may not be predictable from a user's level of expertise. Further complications arise if user preferences change over time. The authors describe a system that begins as a simplified Bayesian network that predicts what a user would like done and then "grows" to increase prediction accuracy. A task in which subjects must find a route through a mine field is used to study the problems that arise with adaptable interfaces.

Colcombe, A. M., Kramer, A. F., Irwin, D. E., & Hahn, S. (2000)
Attentional and oculomotor capture by onset, luminance, and color singletons

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 115–119

> Three experiments investigated whether attentional and oculomotor capture occur only when abrupt onsets that define new objects are used as distracters in a visual search task or whether other salient stimuli also capture attention and the eyes when they do not constitute new objects. The results show that abrupt onsets (new objects) are especially effective in capturing attention and the eyes but that luminance increments that do not constitute a new object capture attention as well. Color singletons do not capture attention unless subjects have experienced the color singleton as a search target in a previous experimental session. Both abrupt onsets and luminance increments elicit reflexive, involuntary saccades, whereas transient color changes do not. These data are discussed in terms of how displays might be designed to aid users in rapidly and accurately extracting needed information. Implications for underlying neuro-anatomical mechanisms, cognition, and aging are discussed.

Colmenarez, A. J. & Huang, T. S. (1996)
Maximum likelihood face detection

Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, 307–311

> In this paper, we present a visual learning approach that uses non-parametric probability estimators. We use entropy analysis over the training set in order to select the features that best represent the pattern class of faces and to set up discrete probability models. These models are tested in the context of face detection via maximum likelihood. Excellent results are reported in terms of the correct-answer, false-alarm trade-off as well as in terms of the computational requirements of the systems.

Colmenarez, A. J., & Huang, T. S. (1998)
Face detection and recognition

In H. Wechsle (Ed.) *Face recognition: From theory to applications* (pp 174–185). New York: Springer

> Two of the most important aspects in the general research framework of face recognition by computer are addressed here: face and facial feature detection and face recognition—or rather, face comparison. The best reported results of the mug shot face recognition problem are obtained with elastic matching via jets. In this approach, the overall face detection, facial feature localization, and face comparison are performed in a single step. This paper describes our research progress toward a different approach for face recognition. On the one hand, we describe a visual learning technique and its application to face detection in complex background and accurate facial feature detection/tracking. On the other

hand, a fast algorithm for two-dimensional template matching is presented, as well as its application to face recognition. Finally, we report an automatic, real-time face recognition system.

Colmenarez, A., Lopez, R., & Huang, T. S. (1997)
Three-D model-based head tracking

Proceedings of the International Society for Optical Engineering, **3024**(pt. 1), 426–434

In this paper, we introduce a new approach to feature-based head tracking and pose estimation. Head tracking and pose estimation find their most important applications in motion analysis for model-based video coding. The proposed algorithm employs an underlying three-dimensional head model, feature-based pose estimation, and texture mapping to produce accurate templates for the feature tracking. In this way, the set of templates used for the matching is constantly updated with the pose changes, which allows the algorithm to track the features over a large range of head motion without error accumulation and loss of precision. Given a rough estimate of the head scale, the initial feature identification is performed automatically and the tracking is successful over a large number of video frames. Computational complexity is also considered with the aim toward creating a real-time end-to-end model-based video coding system.

Darkow, D. J., & Marshak, W. P. (1998)
In search of an objective metric for complex displays

Proceeding of the Human Factors & Ergonomics Society 42nd Annual Meeting, **2**, 1361–1365

Advanced displays for military and other user interaction-intensive systems need objective measures of merit for analyzing the information transfer from the displays to the user. A usable objective metric for display interface designers needs to be succinct, modular, and scalable. The authors have combined the concepts of weighted signal-to-noise ratio and multidimensional correlation to calculate a novel index of display complexity. Preliminary data are presented that support the development of this metric for complex visual, auditory, and mixed auditory and visual displays. Analysis of the human subject data indicates that the coefficients for the algorithm are easily determined. Furthermore, the metric can predict reaction times and accuracy rates for complex displays. This combination of semi-automated reduction of display information and calculation of a single complexity index makes this algorithm a potentially convenient tool for designers of complex display interfaces.

Davis, E., Ntuen, C. A., Perry, A. R., & Marshak, W. P. (2000)
An application of true depth display in visualization of military symbols

Proceedings of the 4th Annual Federated Laboratory Symposium, 135–138

The primary objective of this study was to evaluate the effects of the true depth display (TDD) on human performance during military symbol visualization. The TDD is a device that simultaneously presents two

images in the same visual space, but one image can be manipulated to create differences in depth between the two images. Two experiments were conducted to determine whether depth and clutter affect a person's detection sensitivity. Signal sensitivity is defined as the difference between the probability of detecting a correct signal and the probability of a false alarm in the presence of noise. Results of detection sensitivity are presented here. The major observation was that depth and clutter affect a person's information detection sensitivity.

Dunn, R. S. (1999)
Visualization architecture technology
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 145

The goal of the Crewstation Technology Laboratory is to develop and demonstrate the power of visualization architecture technology (VAT) to depict tactically relevant information during complex operations in a command and control environment. To this end, advanced, three-dimensional stereoscopic display systems must be integrated with high-resolution geo-referenced imagery running on a real-time communications network managed by an executive scenario controller. As a component of VAT, the Force Operational Readiness Combat Effectiveness Simulation (FORCES) controls tactical scenarios illustrating a variety of information visualization concepts. The development of flexible VAT architecture permits the evaluation of information handling and processing and the assessment of decision aids during the design phase of future systems. The ultimate goals include intensifying command situation awareness and increasing the tempo of operations, as well as improving mission planning and control.

Ellis, C. D., & Johnston, D. M. (1999)
Qualitative spatial representation for situational awareness and spatial decision support
In Freksa & Mark (Eds.) *Spatial information theory: Cognitive and computational foundations of geographic information science: COSIT '99.* Berlin: Springer-Verlag

This paper summarizes research on the effectiveness of qualitative spatial representation (QSR) in two-dimensional and three-dimensional displays for improving situation awareness and decision making. The study involved (1) creating spatial query functions based on QSR, which capture knowledge about objects in space; (2) building these query functions into a graphical user interface environment as simulated user-accessible support functions; and (3) testing the utility of these support functions by evaluating the performance of human subjects in solving sets of spatial decision-making and information-retrieval tasks.

Fiebig, C. B. (1999)
Designing experience-centered planning support systems
Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign

> A design methodology known as DAISY (Design Aid for Intelligent Support Systems) is used to develop computer planning support systems that meet the special needs of users at specified levels of experience. In this iterative methodology, the designers observe experts and nonexperts to develop models of the planning tasks and to identify the information and knowledge used by each group. Focusing on differences between the groups, the designers identify specialized system requirements needed to meet the information and display needs of users at a given level of experience. The effectiveness of DAISY was illustrated by its application to the design of the planning support system called Fox, a software application that generates friendly courses of action (FCOAs). Two evaluations showed that Fox significantly increased the range of FCOA options considered by expert users.

Fiebig, C. B., & Hayes, C. C. (1998)
DAISY: A design methodology for experience-centered planning support systems
IEEE International Conference on Systems, Man, and Cybernetics, 1, 920–925

> Designing systems to effectively assist planners in grasping a situation quickly and in making high quality decisions is very difficult, even within a single problem-solving domain. Different types of users have very different needs, and a system designed to assist one group of users may frustrate others who have different amounts of experience. In this paper, we present DAISY, a methodology for developing planning aids. This methodology is intended to enable system designers to identify the system requirements needed to meet the information and display needs of users at a given level of experience before the system is designed. These requirements are identified through user problem-solving studies that define a model of the task, the information requirements, and typical user errors. The DAISY methodology is unique in that it identifies the needs of planners with varying levels of experience and allows these specialized user needs to be incorporated into the software design. Unlike other approaches, DAISY provides concrete methods that are specific to the design of decision support systems for planning. We illustrate the use of this methodology in the design of an intelligent agent and human-computer interface called Fox for the military planning task of generating courses of action. This is a complex and difficult decision-making task in which users make life-and-death decisions while they are under extreme time pressure and overloaded with information.

Fiebig, C., Hayes, C., & Parzen, M. (1997)
Development of expertise in complex domains

1997 *IEEE International Conference on Systems, Man, and Cybernetics, Computational Cybernetics and Simulation*, 3, 2684–2689

> In order to develop effective computer critics, tutors, knowledge acquisition systems, and training strategies, it is necessary to understand how human planners' performance evolves as expertise increases. In this paper, we present two studies of the development of expertise in complex domains: manufacturing planning and software development management planning. Experts in each domain rank ordered the plans created by practitioners at various levels of experience from best to worst quality. We did this to assess whether practitioners really did gain skill with increased experience in both fields or whether experts were "self proclaimed." Next, we analyzed the spoken statements of the practitioners to identify the knowledge and problem-solving strategies they used or lacked. We used these data to model the skill development phases in each domain. These models can be used to develop computer tools and training strategies to help practitioners achieve higher levels of competence.

Fiebig, C., Hayes, C., & Schlabach, J. (1997)
Human-computer interaction issues in a battlefield reasoning system

1997 *IEEE International Conference on Systems, Man, and Cybernetics, Computational Cybernetics and Simulation*, 4, 3204–3209

> In this paper, we present a methodology for developing intelligent computer agents and blackboards to assist planners in grasping the situation quickly and making high quality decisions. We illustrate this methodology in the context of a specific military planning task. Military planners have many cognitive challenges when creating and revising battle plans. At all stages of planning, planners must contend with being overloaded with information, while having little time in which to process it. We illustrate use of the methodology in the design of an intelligent agent and human-computer interface for the planning task of course of action (COA) generation.

Fiebig, C. B., Hayes, C. C., & Winkler, R. P. (1999)
What's new in Fox-GA?

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 9–13

> It is very difficult to design planning assistants that are truly effective in helping planners to create high quality plans quickly. In this paper, we present the results of a series of usability assessments that were conducted to determine how Fox-GA affects military planners' problem-solving behavior and what changes needed to be made in the Fox-GA system to make it a more effective tool.

31

Fiebig, C. B., Schlabach, J., & Hayes, C. C. (1997)
A battlefield reasoning system

Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, (Pt. 1), 15–24

> In the following paper, we introduce an architecture for a battlefield
> reasoning system (BRS), which employs a variety of techniques from the
> fields of human-computer interaction and artificial intelligence. We also
> introduce a course of action generator—the first intelligent agent within
> the overall BRS architecture.

Fiebig-Brodie, C. B., & Hayes, C. C. (2000)
Capturing changes in decision-maker behavior

*Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, **2**,*
1111–1116

> A challenge for the designers of decision support systems (DSSs) is that
> the introduction of a decision aid into a complex setting often generates
> new, unexpected user needs. In this paper, we discuss how the iterative
> application of an experience-centered design methodology called DAISY
> provides concrete methods for modeling these changes and for
> identifying new system requirements caused by the introduction of the
> DSS. We illustrate the iterative use of DAISY in the design, evaluation,
> and modification of Fox, a DSS intended to assist expert military users by
> helping them to generate and evaluate a broad range of plan options

Fiebig-Brodie, C. B., & Hayes, C. C. (2000)
Evaluating the utility of decision support tools to assist in Army tasks

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 37–41

> It is very difficult to build a decision support system to support a wide
> variety of users, even when the users all possess some very narrow skill,
> such as expert friendly course of action (FCOA) planning. A DSS is a tool
> intended to improve performance in decision making, possibly by
> providing a way to organize and interpret problem information, by
> critiquing the users' solutions, or by suggesting plausible solutions. Part
> of the reason that it is difficult to design an effective DSS for many users
> is that experts disagree wildly about what is the "best" solution. In this
> paper, we describe user assessments of Fox, a DSS that generates
> candidate FCOAs for military planners. We found that expert users vary
> greatly in their assessment of what constitutes the best FCOA. However,
> the users appeared to uniformly agree that some categories of FCOAs
> were undesirable. Using this information, we were able to develop or
> redesign guidelines for Fox to best suit this varied group's complex
> needs.

Fijalkiewicz, P. (1999)
An intelligent guidance architecture for definition and preparation of the battlefield

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 147

> IGUANA (Intelligent Guidance and User-Adapted Interaction Agent) is a software application that provides context-sensitive, self-adapting assistance to staff planners as they define and prepare the battlefield using interactive computer controls. The battlefield definition can then be used as input for a course of action generator. IGUANA is distinct from previous intelligent user interfaces in that its guidance rules are not static but evolve, based on its interpretation of data about the current application, the system's hardware, the user, the user's task, and the user's environment. The IGUANA guidance agent architecture can also provide support in the form of debriefings that summarize relevant actions of past users and by providing configuration management suggestions that assist the user in adapting the presentation of information. The IGUANA architecture can also provide decision scripts to enable a user to understand the reasoning behind other users' actions. By providing context-sensitive support, the IGUANA framework enables systems to be developed that improve user understanding of the system and user task performance.

Fijalkiewicz, P., & DeJong, G. (1998)
Cheshire: An intelligent adaptive user interface

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 15–19

> To automatically revise an application's interface, adaptive user interfaces (AUIs) interpret data about a system's user, his or her current task, the available hardware, and the user's environment. AUIs improve the user's performance by providing interface configurations adapted to each individual's needs and preferences. As technological advancements are made and software inevitably becomes more complex, AUIs will provide information filtering by presenting displays best suited for each user. A new approach for AUIs uses explanation-based learning (EBL). The EBL-AUI architecture provides a declarative framework for adaptation that allows for the automatic management of display decisions. Most AUIs currently adapt by using a set of static rules. In contrast, the EBL-AUI system provides a framework that enables its rules to be revised. Cheshire is a preliminary computer system implementation of the EBL-AUI architecture.

George, R., English, J., Borhauer, R., Higley, H., & McCoyd, G. (2000)
Developing a distributed collaborative battle planning system integration of legacy display systems
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 95–99

> In this paper, we report about research efforts in developing an environment for distributed collaborative battle planning. A requirement of this effort was to leverage existing software systems as reusable software components. The combat information processor (CIP) is the most significant legacy component in this collaborative environment. The CIP is a visualization tool that allows a commander to assess battlefield situations, monitor field movements, and plan tactical movements more efficiently. High level software architecture has been developed for the Distributed Collaborative Battle Planning System. Major functional components in the CIP were identified, and object wrappers were used to develop the application programming interface model for the CIP. The data structures of the CIP have been investigated and a data model for CIP reverse engineered. In this paper, we describe the high level architecture and provide the details of the developmental effort. General guidelines in transforming legacy systems to a distributed platform and potential pitfalls in distributed development are discussed.

Gharavi-Alkhansari, M., DeNardo, R., Tenda, Y., & Huang, T. S. (1997)
Resolution enhancement of images using fractal coding
Proceedings of the International Society for Optical Engineering, **3024** (Pt. 2), 1089–1100

> The code generated by fractal coding of a digital image provides a resolution-independent representation of the image, since this code can be decoded to generate a digital image at any resolution. When the image is decoded at a size larger than the original encoded image, image details beyond the resolution of the original image are predicted by assuming local self-similarity in image at different scales. In this paper, we (1) present a formulation of how decoding may be done at a higher resolution, (2) evaluate the accuracy of the predicted details using a frequency analysis of fractally enlarged test images, and (3) propose a method for fractal resolution enhancement without the low-frequency loss of information because of fractal coding.

Gharavi-Alkhansari, M, & Huang, T. S. (1997)
Fractal video coding by matching pursuit
Proceedings of the International Conference on Image Processing, **1**, 157–160

> Fractal image and video coders use redundancies present in different scales of natural images for compression. Motion compensation, on the other hand, is a powerful method for exploiting similarities at the same scale in frames of a video sequence. In this paper, a new method is proposed to take advantage of both inter-scale and intra-scale self-

similarities present in video sequences. A rate distortion optimized orthogonal matching pursuit algorithm is used to seamlessly combine motion compensation and fractal techniques into an efficient video coding algorithm.

Ghelani, D. (July, 1998)
Hand tracking in video using active contour models

Master's thesis, The Pennsylvania State University, State College, PA

Active contour models have attracted considerable interest in recent years. Many kinds of active contours and surfaces as well as energy-minimizing schemes have been presented. One example is a snake, an energy-minimizing spline that is influenced by external forces as well as by image forces that pull it toward features such as lines and edges. Snakes are used in a number of computer vision applications, such as the detection of edges and lines, and in motion tracking and stereo matching. This paper presents an approach in motion (hand) tracking and analysis of deformable objects. The method is based on modeling and extracting the boundary of an object as a generalized active contour model (snake) and then tracking the object boundary in image frames by minimizing the energy function of the contour model. We present an analysis of the contour model (snake) and discuss how the various parameters and forces of the model are selected. The proposed method has been applied to the analysis of a hand tracking experiment. In this method, a snake is used to track a continuous sequence of images captured by video. Results for tracking are presented. Possible failures of the method are also presented.

Goldberg, D. E. (1998)
A meditation on the application of genetic algorithms

Tech. Rep. No. 98003, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory

An argument is presented that genetic algorithms, as search procedures, are not ephemerae, even though they exhibit limitations when shifted from simple, small-scale problems to more complex, real-world ones. Rather than describe successful applications of genetic algorithms, the author accounts for researchers' persistence in employing genetic algorithms by emphasizing the overriding importance of natural selection as an explanatory account of life in the natural environment and the ineffectiveness of traditional optimization and operations research methods.

Goldberg, D. E. (1999)
Using time efficiently: Genetic-evolutionary algorithms and the continuation problem
Tech. Rep. No. 99002, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory

> This paper develops a macro-level theory of efficient time utilization for genetic and evolutionary algorithms. Building on population sizing results that estimate the critical relationship between solution quality and time, the paper considers the trade-off between large populations that converge in a single convergence epoch and smaller populations with multiple epochs. Two models suggest a link between the salience structure of a problem and the appropriate population-time configuration for best efficiency.

Goldberg, D. E., & Pelikan, M. (2000)
Competent and efficient genetic algorithms: Toward computational innovation on the battlefield
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 165

> Decision support tools have been and are being developed to aid the commander in the battlefield, and these tools can be enormously useful. The usual way to view these tools is as information providers or enhancers. An interesting question is whether the commander's search for innovative ways out of a difficult situation can be more directly promoted through the techniques of computational intelligence. Along these lines, the techniques of genetic and evolutionary computation (GEC) are increasingly being used to solve problems across the range of human endeavor. This poster explores (1) some of the recent and stunning progress in achieving genetic evolutionary algorithm (GEA) competence, (2) several techniques of leveraging that competence through a variety of efficiency enhancement methods, and (3) the connection between GEA efficiency and competence on the one hand and human innovation and creativity on the other. The poster concludes by suggesting that GEAs will continue to help us solve difficult problems, quickly, reliably, and accurately, and our greater understanding of human innovation gleaned from GEA practice and theory will help us better design human organizations and institutions.

Goldberg, D. E., & Voessner, S. (1999)
Optimizing global-local search hybrids
Tech. Rep. No. 99001, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory

> This paper develops a framework for optimizing global-local hybrids of search or optimization procedures. The paper starts by idealizing the search problem as a search by a global algorithm G for either (1) acceptable *targets* (solutions that meet a specified criterion) or (2)

basins of attraction that then lead to acceptable targets under a specified local search algorithm *L*. The paper continues by abstracting two sets of parameters: probabilities of successfully hitting targets and basins and time-to-criterion coefficients. With these parameters, equations may be written to account for the total time of search and for the probabilistic success (reliability) in reaching an acceptable solution. Thereafter, optimization problems are formulated in which the division of local versus global search time is optimized so that solution time to acceptable reliability is minimized, or reliability under specified solution time is maximized. A two-basin optimality criterion is derived and applied to important representative problems. Continuations and extensions of the work are suggested, but the theory appears to be immediately useful in better understanding the economy of hybridization.

Goodwin-Johansson, S., Mancusi, J., & Nwankwo, H. (1997)
Applications of tactile interfaces
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 121–130

Future military communication and information systems will give soldiers access to extraordinary amounts of data. The availability of such systems hastens the need for understanding the capacity of human information processing and communication abilities. Heretofore, primarily visual and aural displays have been explored as a means of presenting information. Tactile information, transmitted to the user through the sense of touch and by the user through sensors that detect force, motion, or position, is a communication channel largely unused in legacy military systems, except for computer data entry through keyboards. In this paper, we explore two potential systems that enable communication through tactile interfaces. Two Army user groups have been identified: foot soldiers and computer operators in moving vehicles. Foot soldiers have relatively simple informational needs and are often in situations when traditional tactile interfaces such as keyboards are impractical. Computer operators, on the other hand, are subjected to limited space, whole body vibration and jarring motions, and postural conditions that have implications for tactile interface applications. This paper discusses proposed tactile interfaces for input and output between the soldier and a communication system or computer. The interactions between human factors related to tactile interfaces and the proposed devices are discussed, with the goal of designing interfaces that increase the productivity of the users.

Goodwin-Johansson, S., Mancusi, J., Yadon, L., & Mion, C. (2000)
Flexible input tactile device
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 167

The use of pointing devices has become an integral part of controlling the information displayed on the computer screen. As the use of computers

becomes more integrated into the execution of military operations, there is a need to operate computers in environments that are farther removed from office conditions. Controlling a pointing device inside a military vehicle that is crossing a battlefield can be very difficult because of the vibrations, jolts, and swaying of the vehicle. To improve the manipulation of data on a display, we have proposed a flexible tactile input device that could be attached to the clothing over a soldier's thigh or on other body locations, with a strap to keep the hand in position. This should reduce the relative motion between the hand and the input device because of vehicle motion. This poster discusses the need for the device, and presents the design and fabrication of the second generation prototype.

Goodwin-Johansson, S., Palmer, D., Mancusi, J., Nwankwo, H., Wesler, M., & Marshak, W. (1999)
Tactile interface on a mobile computing platform
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 51–55

> Tactile devices can be used by dismounted soldiers to augment the traditional visual and auditory communication channels. We conducted experiments to investigate the use of an experimental first generation system of tactile devices controlled by a portable computer (DASHER) to convey directional information to the dismounted soldier. Two experiments were performed. The first experiment investigated the ability of a subject to correctly identify which of five spatially separate vibratory tactile devices was actuated. The second experiment investigated the ability of a subject to use vibratory tactile input from five devices to identify 18 different directions. The results of the first experiment indicated that subjects could correctly identify which device was actuated between 82% and 98% of the time for the strong vibration level. The results of the second experiment indicate that if we use combinations of actuators operating at different vibration levels, five actuators are sufficient to communicate to a soldier 18 different directions.

Goodwin-Johansson, S., Yadon, L., Pace, C., & Mion, C. (2001)
Flexible input tactile device
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 101–104

> The use of pointing devices on computers has become an integral part of controlling the information displayed on the computer screen. As the use of computers becomes more integrated into the execution of military operations, there is a need to operate computers in more environments that are farther removed from office conditions. Controlling a pointing device inside a military vehicle that is crossing a battlefield can be very difficult because of the vibrations, jolts, and swaying that the passengers experience. To improve the manipulation of data on a display, we have proposed and fabricated a prototype flexible tactile input device that could be attached to a soldier's clothing over the thigh, with a strap to

keep the hand in position, or on other body locations. This should reduce the relative motion between the hand and the input device, which is attributable to vehicle motion. Some of the testing results from these devices are presented.

Gupta, M. P. (1999)
Reservation-based distributed resource management
Master's thesis, University of Illinois, Urbana-Champaign

> An architecture is described that allows a process to reserve resources on a remote host. The architecture incorporates a resource agent on all hosts involved in a distributed application. These agents are connected and provide for transfer of reservation information among themselves. A request for distributed reservation is made with one of the agents. The request is split according to process locations, and individual components are sent to corresponding agents. The agents in turn interact with various brokers and reserve resources. A broker specializes in management of a single resource in a single end system. The prototype implementation provides reservation for CPU cycles. As brokers for other end system resources are developed, they can be easily incorporated into the architecture.

Hahn, S., & Kramer, A. F. (1998)
Further evidence for the division of attention among noncontiguous locations
Visual Cognition, 5, 217–256

> An investigation was made of the boundary conditions regarding the ability to divide attention among different locations in visual space. In each of five studies, undergraduates (aged 18 to 33 years) performed a same-difference matching test with target letters that were presented on opposite sides of a set of distracter letters. Experiments 1, 2, and 3 provide further support for the proposal that subjects can concurrently attend to noncontiguous locations as long as new objects do not appear between the attended areas. Experiment 4 examined whether the disruption of multiple attentional foci was the result of the capture of attention by new objects per se or by task-irrelevant objects. Multiple attentional foci could be maintained as long as distractor objects did not appear between target locations. Experiment 5 examined whether attention can be divided among noncontiguous locations within as well as between hemifields. Hemifield boundaries did not constrain the subjects' ability to divide attention among different areas of visual space. The results are discussed in terms of the nature of attentional flexibility and putative neuro-anatomical mechanisms that support the ability to split attention among different regions of the visual field.

Han, S., & Wilkins, D. C. (2000)
Efficient computation on minimum error tree Bayesian networks

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 169

> Bayesian networks are an important method of decision making in the presence of uncertainty in artificial intelligence. Exact algorithms for Bayesian inference are computationally intensive for large networks. The inference time grows exponentially with the size of the Bayesian network, which makes the use of large networks impractical in domains that require real-time decision making. This paper reports the first experimental results obtained by the use of minimum error tree decomposition (METD) to increase to speed of Bayesian inference. A learning procedure is described that restructures a Bayesian network as a tree, by the introduction of hidden variables, even when there are errors in the correlation data among the input variables. Experimental results show that speed of computation increases by 3 orders of magnitude for networks with 100 nodes and a connectivity level of two. This allows problems to be solved in 2 seconds, which previously required more than an hour of computing time. In Bayesian networks with a prediction accuracy of 60% to 70%, the use of the minimum METD degraded the prediction accuracy only an additional 10% to 20%.

Harik, G., Cantú-Paz, E., Goldberg, D. E., & Miller, B. L. (1999)
The gambler's ruin problem, genetic algorithms, and the sizing of populations

Evolutionary Computation, 7, 231–253

> A model is presented to predict the convergence quality of genetic algorithms (GAs), based on the size of the population. The model is based on an analogy between selection in GAs and one-dimensional random walks. Using the solution to a classic random-walk problem (the gambler's ruin), the model naturally incorporates previous knowledge about the initial supply of building blocks (BBs) and correct selection of the best BB over its competitors. The result is an equation that relates the size of the population with the desired duality of the solution, as well as the problem size and difficulty. The accuracy of the model is verified with experiments using additively decomposable functions of varying difficulty. The paper demonstrates how to adjust the model to account for noise present in the fitness evaluation and for different tournament sizes.

Harik, G., & Lobo, F. G. (1999)
A parameter-less genetic algorithm

Proceedings of the Genetic and Evolutionary Computation Conference, 258–265

> From the users' point of view, setting the parameters of a genetic algorithm (GA) is far from a trivial task. Moreover, users are typically not interested in population sizes, cross-over probabilities, selection rates, and other GA technicalities. They are interested in solving a problem and

would like to hand the problem to a "black-box" algorithm and simply press a start button. This paper explores the development of a GA that fulfills this requirement by having no parameters whatsoever. The development of the algorithm takes into account several aspects of the theory of GAs, including previous research work on population sizing, the schema theorem, building block mixing, and genetic drift.

Hayes, C. C., & Fiebig-Brodie, C. B. (2000, July)
Community builder: A methodology for designing mixed initiative multi-agent systems

Paper presented at the Sixth International Conference on Intelligent Autonomous Systems (IAS-6), Venice, Italy

> It is difficult for one to develop or adapt and re-use agent-based systems in new task domains. Part of the reason is that each task domain is a little different, requiring modifications of the high-level organization and communications between agents if they are to perform efficiently in the new domain. Much of the work in agent systems focuses on the building blocks of agent systems, such as communication protocol languages, control schemes, and general architectures, and on designing fully automated multi-agent systems, but there is little work on how to organize these building blocks to meet the needs of specific domains in which some of the agents will be human. Community builder is a methodology that assists designers of agent systems in identifying the constraints imposed on the system by the task domain. Our goal in developing community builder is to facilitate faster, more systematic construction of mixed initiative multi-agent decision support systems.

Hayes, C., Penner, R. Ergan, H., Lu, L., Tu, N. Jones, P., Asaro, P., Bargar, R., Chernyshenko, O., Choi, I., Danner, N., Mengshoel, O., Sniezek, J., & Wilkins, D. (2000)
CoRaven: Model-based design of a cognitive tool for real-time intelligence monitoring and analysis

*Proceedings of the 2000 IEEE International Conference on Systems, Man, and Cybernetics, **2**, 1117–1122*

> We describe a model-based design method to develop CoRaven, a decision support tool intended to assist military intelligence analysts in managing and interpreting large quantities of battlefield information. In this method, we use observations of practitioners solving specific tasks in order to understand and model how they use information. We use this model of the task to help identify user needs that the tool must support and to guide usability analyses during initial prototyping. We have found task models to be an important consideration in the decision support tool design process, which can help to constrain the design space and reduce the time required to develop effective decision support tool prototype.

Hayes, C. C., Schlabach, J. L., & Fiebig, C. B. (1998)
FOX-GA: An intelligent planning and decision support tool

Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, **3,** 2454–2459

Fox-GA is described, which is an intelligent planning decision support tool for assisting military intelligence and maneuver battle staff in rapidly generating and assessing battlefield courses of action (COAs). The motivations behind Fox stem from the need to plan and re-plan rapidly to allow users flexibility and control over planning objectives and options. The environment in which plans are executed (the battlefield) is inherently uncertain and rapidly changing, demanding frequent re-planning during execution. To help meet these rapid re-planning needs, we designed Fox to rapidly generate and evaluate a broader variety of high quality COAs faster than military staff could do themselves. Fox then evaluates the COAs and presents only the best few to users, allowing users to reassess those options according to their own judgment and to either edit or select the ones they feel are best. Early evaluations indicate that users explore a wider variety of COAs with Fox than without.

Higley, H. C., & George, R. (2000)
A new approach to building an automated system from incompatible components

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 171

Battlefield planning is a highly distributed process. The primary planning takes place at a high-level command post, with several "what-if" scenarios enacted at lower field commands. Information is usually passed among these levels on paper because most subsystems of a battlefield planning system are not compatible. Each subsystem usually performs one specific function that requires unique hardware and software. This poster describes the evolution of new software engineering tools that have been developed to allow subsystems to be joined into one coherent system.

Hong P., & Huang, T. (1999)
Natural mouse—a novel human computer interface

Proceeding of the 6th International Conferences on Image Processing, **1,** 653–656

Face tracking allows hands-free human-computer interactions. In spite of advances in computer hardware and efficient and robust vision algorithms for tracking, the requirements for effective face and facial tracking in particular situations remains unclear. This paper considers the problems of building a human-computer interface via face tracking and describes an architecture for a novel tool, the *natural mouse.* Natural mouse allows the merging of state-of-the-art face-related techniques and human demand. The advantage of the natural mouse is that people can

dynamically configure it and control it by facial expressions and face motions without the need to wear any accessory equipment. A mouse icon, displayed on the computer monitor, serves as feedback to users of the natural mouse. An immediate application of the natural mouse will be to enable people with hand and speech disabilities to communication with a computer.

Hong, P., Wen, Z., Huang, T. S., & Chan, M. T. (2001)
Speech-driven avatars

Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 119–124

Visual representation of the human is important for a display in collaborative environments. However, in a very dynamic battlefield, a clean and high bandwidth communication channel cannot be guaranteed. A graphics-based human model (i.e., avatar) provides an effective solution. This paper presents three approaches that require only very low bandwidth to drive a remote avatar. The first one is an off-line approach; the second is a real-time speech-driven avatar with a short constant delay. The third approach uses visual cues of speech to drive an avatar and synchronizes the animation sequence with the speech stream.

Huang, T. S. (1997)
Computer vision: Evolution and promise

Proceedings of 5th International Conference on High Technology: Imaging Science and Technology, Evolution and Promise. World Techno Fair in Chiba '96, 13–20

In this paper, I give a somewhat personal and perhaps biased overview of the field of computer vision. First, I define computer vision and give a very brief history of it. Then, I outline some of the reasons why computer vision is a very difficult research field. Finally, I discuss past, present, and future applications of computer vision, concentrating on some examples of future applications that I think are very promising.

Huang, T. S. (1997)
Image processing: Some insights

Proceedings of CERN School of Computing, 17–19

After a brief overview of image science and image processing, I concentrate on the topic of image enhancement, restoration, and reconstruction. I offer three insights: (1) Severely degraded images are very difficult to enhance; (2) the crux of successful image enhancement lies in the use of appropriate *a priori* information; (3) wherever possible, one should try to get good quality images. These are illuminated by examples.

Huang, T., Mehrotra, S., & Ramchandran, K. (1996, March)
Multimedia analysis and retrieval system (MARS) project

Paper presented at the 33rd Annual Clinic on Library Applications of Data Processing: Digital Image Access and Retrieval, Urbana-Champaign, IL

> To address the emerging needs of applications that require access to and retrieval of multimedia objects, we have started a MARS project at the University of Illinois. The project brings together researchers interested in the fields of computer vision, compression, information management, and database systems with the singular goal of developing an effective multimedia database management system. As a first step toward the project, we have designed and implemented an image retrieval system. This paper describes the novel approaches for image segmentation, representation, browsing, and retrieval supported by the developed system. Also described is the direction of future research we are pursuing as part of the MARS project.

Huang T. S., Pavlovic, V. I., & Sharma, R. (1996)
Speech/gesture-based human-computer interface in virtual environments

In L. S. Messing (Ed.), *Integration of gesture in language and speech* (pp 41–58). Wilmington, DE: WIGLS

> Combining machine interpretation of hand gestures and speech can help in achieving the case and naturalness desired for human-computer interaction (HCI). In this paper, investigation of model parameters and analysis of features and their impact on the interpretation of hand gestures are presented in light of the naturalness desired for HCI. Further work that combines advances in computer vision and speech understanding with HCI will be necessary to produce an effective and natural hand gesture interface.

Huang, T. S., Ramchandran, K., Smith, M. J. T, & Farvardin, N. (1999)
Image and video compression: Meeting the Army needs

Joint Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Disp lays & Interactive Displays Consortium, 7–14

> The Army needs compression technologies for multi-spectral and multi-sensor images and video that are high performance, low complexity, scalable, interpretable, and robust to noise. Performance includes not only a satisfactory compression ratio but also good target recognizability and ease of manipulation in the compressed domain. This paper highlights work in data compression undergoing study in the three Army FedLab consortia.

Huang, T., Stroming, J., Kang, Y., & Lopez, R. (1996)
Advances in very low bit rate video coding in North America

IEICE Transactions on Communications, **E79-B**(10), 1425–1433

> Research in very low bit rate video (VLBV) coding has made significant advancements in the past few years. Most recently, the introduction of the MPEG-4 proposal has motivated a wide variety of approaches aimed at

44

achieving a new level of video compression. In this paper, we review progress in VLBV categorized in three main areas: (1) waveform coding, (2) two-dimensional content-based coding, and (3) model-based coding. When appropriate, we also describe proposals to the MPEG-4 committee in each of these areas.

Huang, J., & Zhao, Y. (1997)
Energy-constrained signal subspace method for speech enhancement and recognition
IEEE Signal Processing Letters, 5, 283–285

> An improved signal-subspace-based speech enhancement algorithm is proposed for automatic speech recognition in an additive noise environment. The key idea is to match the short-time energy of the enhanced speech signal to the unbiased estimate of the short-time energy of the clean speech. This technique has proved very effective for improving the estimation of the low-energy segments of continuous speech in low-noise conditions. Experimental results show significant improvement in both the segmental signal-to-noise ratios (SNRs) and the word recognition accuracy of the enhanced speech with SNRs of 10 to 20 dB.

Huang, J., & Zhao, Y. (1997)
A rescaled signal subspace method for speech enhancement and recognition
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 107–120

> In this paper, an improved signal subspace-based speech enhancement algorithm is proposed for automatic speech recognition in an additive noise environment. The underlying principle of the signal subspace algorithm is to decompose the vector space of the noisy signal into a signal-plus-noise subspace and a pure noise subspace. Enhancement is performed by removing the noise subspace and using a linear estimator to estimate the clean speech from the remaining signal-plus-noise subspace. A rescaling method is developed to adjust the short-time energy of the estimated signal, referred to as the rescaled signal subspace (RSS) method. It is shown that the RSS method is very useful for improving the estimation of the unvoiced and transition segments of continuous speech. As a result, this method improved the recognition accuracy significantly in low SNR conditions. Furthermore, a signal subspace rotation algorithm is combined with the RSS method, resulting in an improved method of speech recognition called the signal subspace rotation (SSR) method. The key idea is to rotate the signal subspace basis vectors so that better estimation can be made for the low-energy signals in a new subspace. The performances of the algorithms were evaluated with the TIMIT database in the SNR conditions of 5 dB, 10 dB, and 20 dB. We found that the SNR improvements with the RSS and SSR methods

were 2.3 to 6.8 dB. The automatic recognition of the enhanced continuous speech was performed and evaluated, and we found that the RSS and the SSR methods helped to increase the recognition accuracy over the baseline by 11.7% to 88.2% and 12.7% to 92.2%, respectively.

Irwin, D. E., Colcombe, A. M., Kramer, A. F., & Hahn, S. (2000)
Attentional and oculomotor capture by onset, luminance, and color singletons
Vision Research, **40**, 1443–1458

In three experiments, we investigated whether attention and oculomotor capture occur only when abrupt onsets that define new objects are used as distracters in a visual search task or whether other salient stimuli also capture attention and the eyes even when they do not constitute new objects. The results showed that abrupt onsets (new objects) are especially effective in capturing attention and the eyes, but that luminance increments that do not accompany the appearance of new objects capture attention as well. Color singletons do not capture attention unless subjects have experienced the color singleton as a search target in a previous experimental session. Both abrupt onsets and luminance increments elicit reflexive, involuntary saccades whereas transient color changes do not. Implications for theories of attentional capture are discussed.

Iskarous, K. (1999)
Patterns of tongue movement
Proceedings of the 14th International Congress of Phonetic Sciences, 429

This paper discusses the pivot pattern of tongue movement. In this pattern, there seems to be a point in the vocal tract where there is no motion, but there is motion at points of the vocal tract anterior and posterior to the pivot point. Based on tongue edge tracings of frames from ultrasound and x-ray dynamic imaging of the vocal tract, I show that the pivot pattern is used in a variety of sequences, and I discuss the possible causes of the pattern.

Iskarous, K., Baxter, D., Cha, J.-Y., & Morgan, J. L. (1997)
The temporal coordination of gesture and speech
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 1), 35–40

In this paper, we present the results of experiments on the synchronization of pointing gestures and speech. Evidence is presented to show that there is much regularity in the way that pointing gestures are aligned on a small temporal scale with the syntactic boundaries of the phrases that they accompany. Furthermore, it is shown that the alignment of pointing gestures to syntactic domains is sensitive to prosodic effects.

Iskarous, K., & Morgan, J. (1999)
Speech synthesis in a virtual environment

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 149

> A method is described that increases the intelligibility of synthesized
> speech by focusing on the synthesis of stop consonants such as *t, d, k, g,*
> and *n,* which occur frequently enough to hinder understanding of
> synthesized speech. As a solution to this problem, tongue movement
> produced during consonant-vowel frequency transitions is modeled by a
> cubic bezier spline curve whose shape is specified completely by four
> control points. Complex tongue motion during a transition is modeled by
> the movement of only two of these four points, which can be represented
> by a change in a very small number of parameters sampled at 5 to 8
> points. This is an improvement over current systems, which synthesize
> speech by transitioning between concatenated speech sounds by linear or
> higher order frequency interpolation.

Iskarous, K., & Morgan, J. (2000)
Direct modeling of contextual dynamics in stochastic speech recognition

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 73–76

> In this paper, we present a new method for extracting dynamic
> information from the speech signal. This method is based on extracting
> dynamic extractors from a reconstruction of the dynamical system's
> phase space. We then summarize the performance results obtained from
> the new system as implemented in a hidden Markov model recognizer.
> This is the final component in a larger speech recognition system, which
> includes a high-level grammar and a gesture recognizer.

Isarous, K., & Morgan, J. (2001)
Interfacing a speech recognizer and an articulatory synthesizer

Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 73–77

> In this paper, we present recent advances from our laboratory in speech
> production modeling and show the implications of these advances for a
> speech synthesizer and a speech recognizer previously developed for the
> Federated Laboratory project. We also present a novel approach for
> interfacing the synthesizer and recognizer for the purpose of separate
> modeling of different sources of speech variation.

Iskarous, K., Morgan, J., & Cha, J-Y. (1998)
Syntactic and prosodic information in a speech and gesture recognition system
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 11–14

> To enable natural human-computer interaction in a virtual environment, information has to be captured from a number of human communication channels including speech, gesture, gaze, and facial expression. Furthermore, the information from different channels has to be aligned and correlated in order to obtain the overall meaning of the communication act. This paper focuses on relating and aligning the information from the speech and gesture signals. It will be shown that speech-gesture alignment is not a trivial problem and that syntactic and prosodic information is key to the alignment. We then present the architecture of an adaptive hidden Markov model-based speech and gesture recognition system that incorporates the prosodic and syntactic alignment constraints.

Jog, K. (1998)
Stereoscopic calibration of a see-through head-mounted display
Unpublished master's thesis, Pennsylvania State University, University Park, PA

> Abstract not available.

Johnston, D. M., & Ellis, C. (1997)
Representation and computation of spatial relations in the context of situational awareness
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 1), 79–90

> Spatial information may be described as being composed of objects and relations between objects. Specific to spatial objects are definitions of geographic locations. Relations define the nature of interactions between objects. Fundamental spatial relations include *disjoint, meet, contained in, contains*. Extended spatial relations may be defined to include such notions as *in sight of* or *within distance of*. Typically, spatial relations are described with topologic, directional, or metric representations. Human operators extensively employ spatial relations in comprehending geographic space. Most computational environments, however, give limited support for queries using terms related to spatial relations, and most formal models have severe operational constants on them, including limitations to 2-D space or isotropic environments. We summarize the current state of theories and methods of representation of spatial relations and propose experiments to determine the effectiveness of current models of spatial relations in aiding humans in SA activities in the context of different visualization environments.

Johnston, D. M., Ellis, C. D. (1999)
The effectiveness of qualitative spatial representation in supporting spatial awareness and decision making
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 71–75

> This paper summarizes elements of research about the effectiveness of using qualitative spatial representation (QSR) in 2-D and 3-D display modes to determine its usefulness for spatial awareness and decision making. The study involved (1) creating spatial query functions based on QSR that capture knowledge about objects in space; (2) building these query functions into a GUI environment as simulated user-accessible support functions; and (3) testing the utility of these support functions by evaluating the performance of human subjects in solving sets of spatial decision-making and information retrieval tasks.

Jojic, N., Gu, J., Shen, H. C., & Huang, T. S. (1998)
3-D reconstruction of multipart self-occluding objects
In R. Chin & T. C. Pong (Eds.) *Lecture Notes in Computer Science* (pp II-455–II-462). Springer: New York

> In this paper, we present a method for reconstructing multi-part objects from several arbitrary views via deformable super-quadrics as models of the object's parts. Two visual cues are used: occluding contours and stereo (possibly aided by projected patterns). The object can be relatively complex and can exhibit numerous self-occlusions from some or all views. Our preliminary experiments on a human body and a tailor's mannequin show that the reconstruction is more complete than in purely stereo or structured light-based methods and more precise than the reconstruction from occluding contours only.

Jojic, N., & Huang, T. S. (1998)
On analysis of cloth drape range data
In R. Chin & T. C. Pong (Eds.) *Lecture Notes in Computer Science* (pp II-463–II-470). Springer: New York

> In this paper, we present an algorithm for analyzing the range data of cloth drapes. The goal is the estimation of parameters for modeling and the geometry of the underlying object. In an analysis-by-synthesis manner, the algorithm compares the drape of the model with the range data and searches for the best fit. It can be applied to any physics-based cloth model. The motivating application is fashion design using computer-aided design (CAD) systems, but the ability of the algorithm to estimate the shape of the object supporting the scanned cloth indicates the possibility of using cloth models to overcome problems in human tracking algorithms caused by clothing.

Jones, P. M., Hayes, C. C., Fiebig, C., Dunmire, C. (1998)
Cooperative problem solving: A cognitive engineering and distributed cognition view
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 125–129

> Cooperative problem solving is a fundamental part of battlefield operations. Our particular interest is to study cooperative problem solving in a variety of battlefield contexts, with particular emphasis on distributed collaborative planning. In this paper, we propose steps toward an integrated theory and methodology for cooperative problem solving, with a brief example drawn from our experiences at Prairie Warrior '97.

Jones, P. M., Hayes, C. C., Wilkins, D. C., Bargar, R., Sniezek, J., Asaro, P., Mengshoel, O., Kessler, D., Lucenti, M., Choi, I., Tu, N., & Schlabach, J. (1998)
CoRAVEN: Modeling and design of a multimedia intelligent infrastructure for collaborative intelligence analysis
Proceedings of the 1998 IEEE International Conference on Systems, Man, and Cybernetics, **1**, 914–919

> Intelligence analysis is one of the major functions performed by an Army staff in battlefield management. In particular, intelligence analysts develop intelligence requirements based on the commander's information requirements, develop a collection plan, and then monitor messages from the battlefield with respect to the commander's information requirements. The goal of the CoRAVEN project is to develop an intelligent collaborative multimedia system to support intelligence analysts. Key ingredients of our design approach include (1) significant knowledge engineering activities with domain experts, (2) representation of an explicit model of reasoning and activity to drive design, (3) the use of Bayesian belief networks as a way to structure inferences that relate observable data to the commander's information requirements, (4) collaborative graphical user interfaces to provide flexible support for the multiple tasks in which analysts are engaged, (5) sonification of data streams and alarms to support enhanced situation awareness, (6) detailed psychological studies of reasoning and judgment during uncertainty, and (7) iterative prototyping of candidate designs with domain experts. This paper presents our recent progress on all these fronts.

Jones, P. M., Wilkins, D. C., Bargar, R., Sniezek, J., Asaro, P., Danner, N., Eychaner, J., Chernyshenko, S., Schrah, G., Hayes, C., Tu, N., Ergan, H., Lu, L. (2000)
CoRaven: Knowledge-based support for intelligence analysis
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 89–93

> Intelligence analysis is one of the major functions performed by Army staff in battlefield management. This paper reports about a formative evaluation of the first CoRaven prototype and the redesign of CoRaven

based on that evaluation. The CoRaven project's goal is to develop a collaborative multimedia intelligence tool to support intelligence analysis. Key ingredients of our design approach include (1) significant knowledge engineering and iterative prototyping activities with domain experts, (2) task-specific graphical user interfaces that allow multiple ways of viewing battlefield information, (3) Bayesian belief networks to model reasoning on battlefield information, (4) use of sound (sonification) as an additional channel through which to communicate complex data, (5) collaboration technologies, and (6) psychological studies of reasoning and judgment during uncertainty.

Kettebekov, S., & Sharma, R. (2000)
Understanding gestures in multimodal human computer interaction
International Journal on Artificial Intelligence Tools, 9, 205–223

> Because of the advances in recent years in computer vision research, free-hand gestures have been explored as a means of HCI. Gestures in combination with speech can be an important step toward natural, multimodal HCI. However, interpretation of gestures in a multimodal setting can be a particularly challenging problem. We propose an approach for studying multimodal HCI in the context of a computerized map. An implemented test bed allows us to conduct user studies and address issues toward understanding hand gestures in a multimodal computer interface. Absence of an adequate gesture classification in HCI makes gesture interpretation difficult. We formalize a method for "bootstrapping" the interpretation process by a semantic classification of gesture primitives in an HCI context. We distinguish two main categories of gesture classes, based on their spatio-temporal deixis. Results of user studies revealed that gesture primitives, originally extracted from weather map narration, form patterns of co-occurrence with speech parts in association with their meaning in a visual display control system. The results of these studies indicated two levels of gesture meaning: individual stroke and motion complex. These findings define a direction in approaching interpretation in natural gesture-speech interfaces.

Kettebekov, S., & Sharma, R. (2001)
Deriving syntax of gesture/speech for display control
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 95–100

> Gestures in combination with speech can be an important step toward natural, multimodal human-computer interaction (HCI). However, inclusion of non-predefined gestures into a multimodal setting can be a particularly challenging problem. In this paper, we propose a structured approach for studying multimodal language in the context of display control. We consider the systematic analysis of gestures, starting from observable primitives to their semantics. We present a computational framework for gesture-speech integration, which was used to develop an

interactive test bed (iMAP). The results of the studies of the test bed help us understand gesture-speech integration in the context of HCI.

Konrad, C. M., Kramer, A. F., Watson, S. E., & Weber, T. A. (1996)
A comparison of sequential and spatial displays in a complex monitoring task
Human Factors, 38, 464–483

> A sequential display was compared with a more conventional spatial display as 24 college students (aged 17 to 32 years) monitored dynamically changing sets of three-digit numbers and responded to occasional target stimuli. In an effort to equate the stimulus-response compatibility of the two displays, subjects responded to the targets with a chord keyboard in Exp 1 and vocally in Exp 2. The influence of display duration on performance was examined with the sequential and spatial formats by presenting stimuli at durations of 400, 800, and 1200 milliseconds (ms). The influence of practice on performance with the sequential and spatial displays was also investigated. Subjects responded to targets more quickly in the sequential than in the spatial displays at each of the three presentation durations and across more than 2,000 practice trials. Accuracy was influenced by the display presentation duration. Accuracy was higher for the sequential than for the spatial display at the 800-ms stimulus presentation duration in Exp 1 and at the 800- and 1200-ms presentation durations in Exp 2. Results are discussed in terms of the potential utility of sequential displays for complex, real-world systems.

Kothari, J., Grossman, E., & Mehrotra, S. (1998)
Neighborhoods: A framework for enabling web-based synchronous collaboration and hierarchical navigation
Proceedings of the Thirtieth Hawaii International Conference on System Sciences, 1, 666–675

> The World Wide Web (WWW) is an extremely effective mechanism for sharing information throughout the world via a web of links. These links allow anyone with a connection to the Internet to unearth large amounts of information about multitudes of topics. However, access to this information is asynchronous, with no way for users to interact with each other in real time. We have developed a protocol called "Neighborhoods" to support synchronous interaction among users. By grouping related documents together, we can create a virtual neighborhood where WWW users can meet to find and contact others with mutual interests. Neighborhoods are based in underlying protocols for creating collections of documents and for establishing collaborative sessions. The collection protocols allow material on the web to be organized into categories and subcategories, providing context and a readable graphical representation of a set of related documents. The Neighborhoods protocol adds to this a method for specifying and establishing new collaborative sessions, as well as locating existing ones.

Kramer, A. F., Hahn, S., Irwin, D. E., & Theeuwes, J. (1998)
Attentional capture and oculomotor control
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 91–95

> People make rapid eye movements to examine the world around them. Before an eye movement is made, attention is covertly shifted to the location of the object of interest, and the eye typically will land at the position where attention is directed. Here we report that a goal-directed eye movement toward an object is disrupted by the sudden appearance of a task-irrelevant object. In many instances, before the eye reached the target, it started moving in the direction of the new object. The eye often landed for a very short period of time (25 to 150 milliseconds) near the new object. The results suggest parallel programming of two saccades: one voluntary goal-directed eye movement toward the color target and one stimulus-driven eye movement reflexively elicited by the appearance of the new object. Neuro-anatomic structures responsible for parallel programming of saccades are discussed as is the implication of this research for the presentation of information on complex displays.

Kramer, A. F., Hahn, S., Irwin, D. E., & Theeuwes, J. (1999)
Attentional capture and aging: Implications for visual search performance and oculomotor control
Psychology and Aging, **14**, 135–154

> Two studies were performed that examined potential age-related differences in attentional capture. Subjects were instructed to move their eyes as quickly as possible to a color singleton target and to identify a small letter located inside it. In half of the trials, a new stimulus (i.e., a sudden onset) appeared simultaneously with the presentation of the color singleton target. The onset was always a task-irrelevant distractor. Response times were lengthened, for both young and old adults, whenever an onset distractor appeared, despite the fact that subjects reported being unaware of the appearance of the abrupt onset. Eye-scan strategies were also disrupted by the appearance of the onset distractors. In about 40% of the trials during which an onset appeared, subjects made an eye movement to the task-irrelevant onset before moving their eyes to the target. Fixations close to the onset were very brief, suggesting parallel programming of a reflexive eye movement to the onset and goal-directed eye movement to the target. These data are discussed in terms of age-related sparing of the attentional and oculomotor processes that underlie the phenomenon of attentional capture.

Kramer, A. F., Hahn, S., Irwin, D. E., & Theeuwes, J. (2000)
Age differences in the control of looking behavior: Do you know where your eyes have been?
Psychological Science, **11**, 210–217

Previous research has shown that during visual search, young and old adults' eye movements are equivalently influenced by the appearance of task-irrelevant abrupt onsets. The finding of age-equivalent oculomotor capture is quite surprising in light of the abundant research suggesting that older adults exhibit poorer inhibitory control than young adults on a variety of different tasks. In the present study, the authors examined the hypothesis that oculomotor capture is age invariant when subjects' awareness of the appearance of task-irrelevant onsets is low but that older adults will have more difficulty than young adults in inhibiting reflexive eye movements to task-irrelevant onsets when awareness of these objects is high. Nineteen old (67 to 75 years) and 19 young (18 to 25 years) adults participated in the study. Subjects' awareness of task-irrelevant onsets was varied by, in one condition, making onset equi-luminent to other stimuli in the display and in the other condition by making onset brighter than the other stimuli. Results were consistent with the level-of-awareness hypothesis. Young and old adults showed equivalent patterns of oculomotor capture with equi-luminant onsets, but older adults misdirected their eyes to bright onsets more often than young adults did. These findings are discussed in terms of their implications for the nature of inhibitory processes that underlie eye movements and visual attention.

Kramer, A. F., Larish, J. L., Weber, T. A., & Bardell, L. (1999)
Training for executive control: Task coordination strategies and aging
In D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII: Cognitive regulation of performance: Interaction of theory and application.* (pp 617–652). The MIT Press: Cambridge, MA

The authors studied the ability to successfully coordinate the performance of multiple tasks as a function of two multi-task training strategies, variable priority (VP) training and fixed priority (FP) training. The acquisition, retention, and transfer of task coordination skills was investigated in adults, both young (aged 18 to 29 years) and old (60 to 75 years). After training in two tasks (a canceling and a tracking task), each of which possessed both repeating and random sequences, the authors asked subjects to perform several novel versions of the two tasks in an effort to evaluate learning of the repeated patterns in the single- and dual-task conditions. The authors then had the subjects perform two novel tasks in an effort to examine the generalizability of task coordination skills acquired during VP and FP training. Finally, retention of the original training tasks was assessed in single- and dual-task conditions 45 to 60 days after the training intervention. Results indicated that subjects who trained with the VP procedure learned the training tasks more

quickly and exhibited a higher level of mastery of the tasks than did subjects trained with the FP technique. Furthermore, the decrement in dual-task performance usually found in older adults (and observed before training in the older adults in this study) was substantially reduced for the VP-trained subjects but not for the FP-trained subjects. Finally, subjects trained with the VP procedure exhibited better transfer to novel tasks as well as higher levels of retention than did FP-trained subjects.

Kramer, A. F., & Weber, T. A. (1999)
Applications of psychophysiological techniques to human factors
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 85–89

This paper provides a brief overview and critical review of two different potential applications of psychophysiological techniques to important issues in human factors: the assessment of fluctuations in alertness and the use of psychophysiological measures in on-line adaptive algorithms. The advantages and disadvantages of using psychophysiological measures in these domains are described, and the potential for further development of psychophysiologically based assessment of mental processing and operator state is discussed.

Kramer, A. F., & Weber, T. A. (1999)
Object-based attentional selection and aging
Psychology and Aging, **14**, 99–107

Two studies were conducted that examined potential age-related differences in object-based attentional selection. In both studies, subjects were briefly presented with pairs of wrenches and asked to make one response if two target properties (i.e., an open end and hexagonal end) were present and another response if only a single target property was present in the display. The critical manipulation was whether the target properties were present on one wrench or distributed between two wrenches. Space-based models of selective attention predict no difference in performance between these conditions. However, object-based attentional selection models predict better performance when both target properties appear on a single object. The results from both studies were consistent with object-based models of attentional selection. Furthermore, both young and old adults showed similar performance effects, suggesting that object-based attentional selection is insensitive to normal aging.

Kramer, A. F., & Weber, T. (2000)
Applications of psychophysiology to human factors
in J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.) *Handbook of psychophysiology* (2nd ed., pp 794–814). New York: Cambridge University Press

We discuss current problems in human factors amenable to study by psychophysiological methods. First, we identify (a) current topics of

interest in human factors, (b) the criteria that psychophysiological measures must meet to be useful in human factors applications, and (c) the history of psychophysiological methods in human factors. Next, we turn to three specific applications for psychophysiological measures: vigilance decrements, alertness, and mental workload, and we conclude with a discussion about whether cognitive constructs are each indexed by a unique psychophysiological measure.

Kramer, A. F., Weber, T. A., & Watson, S. E. (1997)
Object-based attentional selection—Grouped arrays or spatially invariant representations?: Comment on Vecera and Farah (1994)
Journal of Experimental Psychology: General, **126,** 3–13

> S. P. Vecera and M. J. Farah addressed the issue of whether visual attention selects objects or locations. They obtained data that they interpreted as evidence for attentional selection of objects from an internal spatially invariant representation. Kramer, Weber, and Watson question this interpretation on both theoretical and empirical grounds. First, the authors suggest that there are other interpretations of the Vecera and Farah data that are consistent with location-mediated selection of objects. Second, they provide data, using the displays employed by Vercera and Farah along with a post-display probe technique, suggesting that attention is directed to the locations of the target objects. The implications of the results for space- and object-based attentional selection are discussed.

Lazaridis, I., & Mehrotra, S. (2001)
Incorporating aggregate queries in interactive visualization
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays and Interactive Displays Consortium, 125–130

> This paper discusses a new data structure (i.e., the multi-resolution aggregate tree [MRA-tree]) that can be used to give approximate answers to spatial aggregate queries. Spatial aggregate queries involve asking for the value of some aggregate function for a specific region of space. Examples are "What is the average wind velocity for the next 100 miles of my flight path at a 1-mile resolution" or "What is the total number of vehicles within 10 miles of my position." Our technique handles all the common types of structured query language aggregates (MIN, MAX, SUM, COUNT, AVG). We specify how to estimate the aggregate, using the nodes of the MRA-tree, and how to give tight 100% intervals of confidence on the actual value of the aggregate. We also propose a tree-traversal strategy that reduces the error as more tree nodes are explored. Using an MRA-quadtree in experiments employing both real and synthetic data sets, we have shown the validity of our approach for fast computation of spatial aggregates for even exact answering, indicating that our method can be used in performance-sensitive virtual geographic information systems.

56

Li, Y., & Zhao, Y. (1998)
Recognizing emotions in speech using short-term and long-term features

Proceedings of the 5th International Conference on Spoken Language Processing, **6**, 2255–2258

> The acoustic characteristics of speech are influenced by speakers' emotional status. In this study, we attempted to recognize the emotional status of individual speakers by using speech features extracted from short-time analysis frames as well as speech features representing entire utterances. Principal component analysis was used to analyze the importance of individual features in representing emotional categories. Three classification methods were used, including vector quantization, artificial neural networks, and a Gaussian mixture density model. Classifications using short-term features only, long-term features only, and both short-term and long-term features were conducted. The best recognition performance (of 62% accuracy) was achieved when the Gaussian mixture density method was used with both short-term and long-term features.

Lin, J., Wu, Y, & Huang, T. S. (2001)
Modeling the natural hand motion constraints

Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 105–110

> Capturing hand motion is one of the most important steps in constructing a gesture interface. Many current approaches to this task generally involve a formidable nonlinear optimization problem in a large search space. However, if one takes into account the constraints on hand motion (fingers can be curled in one direction only) a significant reduction in the size and dimensionality of the search space can be achieved. In this paper, we propose a learning approach to model the hand configuration space directly from motion data collected from a CyberGlove. We eliminate the redundancy of the feasible configuration space by finding a more compact and efficient representation of the original space in a lower dimensional subspace. Based on the linear behavior observed in this subspace, finger configurations are modeled by the union of these linear manifolds. This motion-constraint model enables improved and efficient motion capturing from video input. Several experiments show how we capture articulated hand motion by taking advantage of our proposed model.

Lopez, R., Colmenarez, A., & Huang, T. S. (1997)
Vision-based head and facial feature tracking

Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 73–84

> In the following paper, we introduce an algorithm for automatic head tracking, using a model-based approach. The input to the system is a two-dimensional video sequence of a person's head and shoulders, and the

output consists of the trajectories of salient facial features, as well as an estimate of the three-dimensional (3-D) motion of the head. Issues such as localization accuracy and error accumulation are overcome by using an underlying 3-D model to complete optimal templates for each video frame for use in the feature-tracking module. The algorithm has been tested on synthetic and real sequences and is shown to produce accurate results for more than 100 frames at approximately five frames per second.

Loschky, L. C., & McConkie, G. W. (1999)
Gaze contingent displays: Maximizing display bandwidth efficiency

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 79–83

> One way to economize bandwidth in single-user head-mounted displays is to put high-resolution information only where the user is currently looking. This paper describes a series of six research projects investigating spatial, resolutional, and temporal parameters affecting perception and performance in eye-contingent multi-resolutional displays. Based on the results of these projects, suggestions are made for the design of eye-contingent multi-resolutional displays.

Loschky, L. C., & McConkie, G. W. (2000)
User performance with gaze-contingent multiresolutional displays

In A. T. Duchowski (Ed.), *Eye-tracking research & applications symposium* 2000 (pp 97–103). New York: Association for Computing Machinery

> One way to economize bandwidth in single-user HMDs is to put high-resolution information only where the user is currently looking. This paper summarizes results form a series of six studies investigating spatial, resolutional, and temporal parameters affecting perception and performance in such eye-contingent multi-resolutional displays. Based on the results of these studies, suggestions are made for the design of eye-contingent multi-resolutional displays.

Loschky, L. C., McConkie, G. W., Yang, J., & Miller M. E. (2001)
Perceptual effects of a gaze-contingent multi-resolution display based on a model of visual sensitivity

Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 53–58

> Many interactive single-user image display applications have prohibitively large bandwidth requirements. However, bandwidth can be greatly reduced by using gaze-contingent multi-resolution displays (GCMRDs) that put high-resolution only at the center of vision, based on eye position. A study is described in which photographic GCMRD images were filtered as a function of contrast, spatial frequency, and retinal eccentricity on the basis of a model of visual sensitivity. This model has previously been tested with sinusoidal grating patches. The current study measured viewers' image quality judgments and their eye movement parameters and found that photographic images filtered at a

level predicted to be at or below perceptual threshold produced results statistically indistinguishable from those of a full high-resolution display.

Ma, J. & Ahuja, N. (1998)
Dense shape and motion from region correspondences by factorization
Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 219–224

> In this paper, we propose an algorithm for estimating dense shape and motion of dynamic piecewise planar scenes from region correspondences via factorization. Region correspondences are used since they are easier to establish and more reliable than either line or point correspondences. The image measurements required are the centroid and area for each region. We use singular value decomposition to find the basis of range space of the motion, shape, and surface normal matrices. By imposing model constraints, we can recover motion, shape, and surface normal only from region correspondences.

Ma, J., & Ahuja, N. (1999)
3-D reconstruction from video sequences
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 151

> A process is described to estimate three-dimensional (3-D) structure from two-dimensional video sequences. In contrast to existing methods that use only pixel- or line-based features, the process presented here was a multi-feature-matching algorithm. Image frames are independently segmented at multiple scales, and salient regions are identified across successive video frames, based on characteristics such as region area, moments, intensity values, shape compactness, and adjacency. The 3-D motion and structure of these matched regions are estimated from the established correspondences with a region-based structure-from-motion algorithm. In a second step, the 3-D estimates are used to guide pixel-level matching of the unmatched areas. Candidates for pixel matches are selected in part on the basis of the 3-D motion and structure estimates, and matching is performed in terms of intensity, "edgeness," and "cornerness." Finally, the 3-D structure for each pixel is calculated. From matches of the first three frames, a trilinear tensor can be recovered, which describes the relations between pixels in three images and can be used to predict locations of pixels in subsequent frames. The trilinear tensor provides a general warping function between the pixels in different frames and is used as a measure of confidence for matching in subsequent frames.

Ma, J., & Ahuja, N. (2000)
Region-based motion grouping for augmented reality

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 127–129

In video-based augmented reality, motion grouping is an enabling
technique that aims to break a scene into its most prominent moving
groups that correspond to different moving objects or to objects at
different depths. In this paper, we propose a region-based motion-
grouping algorithm to overcome the difficulties encountered in existing
approaches. First, multi-scale image segmentation is performed on
individual images. Then, the acquired regions are matched via an
eigenspace region-matching algorithm; two-dimensional affine motion
parameters are estimated for each region. The regions are treated as
nodes in a weighted graph, with the weights determined by the
differences of motion. In order to separate the graph into sub-graphs
corresponding to different moving objects, a generalized eigenvalue
system was solved with eigenvectors being the indicators of optimum
partition. The eigenvector with the second smallest eigenvalue is used to
bipartition the graph by finding the splitting point that minimizes an
error measure. Finally, the procedure is performed recursively until there
are no independent moving objects in the scene. Examples of motion
segmentation are presented.

Ma, J., & Ahuja, N. (2000)
Region correspondence by global configuration matching and progressive
Delaunay triangulation

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, **2,**
637–656

In this paper, we present a novel algorithm for establishing region
correspondences across images by first matching global region
configuration and then propagating the matches locally, constrained by
Delaunay triangulation. We exploit a global configuration constraint,
which has not been explicitly used in existing matching algorithms. The
proposed algorithm consists of two stages. In the first, stable regions are
matched by enforcing the global configuration constraint. This yields a set
of global matches corresponding to stable regions distributed over the
images. In the second stage, these matches are used to guide the matching
of the remaining unmatched regions in the intervening spaces. This is
done by enforcing local positioning constraints, which start with the
Delaunay triangulation defined by the global matches, followed by
progressive Delaunay triangulation for local matching. Experiments on
both stereo and motion images are presented to show the effectiveness of
the proposed algorithm.

Marshak, W. P. (1997)

Identifying research areas for the digitization of the battlefield

Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, (Pt. 1), 1–14

>An assessment of research needs was conducted as part of initiating the
>Advanced Displays and Interactive Displays consortium research on
>digitization of the battlefield. Sources for this assessment included an
>extensive search of Army World Wide Web sites, Army publications, site
>visits, and interviews with digitally experienced soldiers. Seven critical
>display and control research needs are identified: computing while "on
>the move," creating the "big picture," collaborating through common
>view and commander's intent, having a common soldier interface,
>integrating legacy systems, assessing bandwidth effects on displays and
>controls, and developing interface evaluation methods. Although most of
>the critical needs are already being studied, the consortium's research
>plans will be adapted to ensure that all the Army's critical needs are met.

Marshak, W. P., & Darkow, D. J. (1998)

Objective measurement of display formats: Multi-dimensional and
multimodal user perception models

Proceedings of the IEEE 1998 International Conference on Image Processing, 2, 505–509

>Comparing the effectiveness of display formats, especially displays set in
>different sensory modalities and containing complex combinations of
>dimensions, can be like comparing the proverbial apples and oranges.
>Dissimilar displays can be compared if a "unit-less" dimension can be
>found that describes how well critical information is expressed, compared
>to other information contained in the display. The signal-to-noise ratio
>(SNR) is such a measure. Fourier power spectra can be computed for
>energy imparted by the display of critical information (signals) and the
>remainder of the display (noise). By computing SNRs for each feature
>channel (modality or dimension), one can obtain complex SNRs to
>describe the salience of the signal. Also considered is the similarity of
>signal and noise as expressed in the Pearson product-moment correlation
>coefficient. Computational examples of such display SNRs are presented
>and discussed.

Marshak, W. P., & Darkow, D. J. (1998)

Prototype depth-separated coincident transparent (true depth) display

Proceeding of the 42nd Annual Meeting of the Human Factors & Ergonomics Society, 2, 1151

>Getting the "big picture" from computer displays is a critical problem for
>user interface designers. Traditional solutions layer information on a
>single display or make multiple displays available simultaneously. These
>strategies fragment the information and require the user to integrate
>information across displays. A new display strategy being developed
>uses depth-separated coincidental transparent displays that we call "true
>depth displays" (TDDs). TDD employs two display surfaces in the same

visual space but separated in depth. Users may read either surface by refocusing their eyes or by focusing between the displays to see both. Display formats can be organized to exploit their spatial coincidence, making integration across displays easy. Information density can be increased without the debilitating effects of clutter. A compact hardware prototype of the TDD was shown along with a variety of format examples to demonstrate the capabilities of this new display technology interface.

Marshak, W. P., Darkow, D. J., Wesler, M. Mc., & Fix, E. L. (2000)
Objective measurement of complex multi-modal and multi-dimensional display formats: A common metric for predicting format effectiveness
Proceedings of SPIE's International Society for Optical Engineering, **4022**, 136–145

Computer-display designers have more sensory modes and more dimensions within sensory modes to encode information than ever before. This elaboration of information presentation has made measurement of display format effectiveness and prediction of a user's performance with the display extremely difficult. A multivariate method has been devised that isolates critical display information, physically measures its signal strength, and compares it with other elements of the display that act like background noise. This method, which we call the *common metric,* relates signal-to-noise ratios (SNRs) within each stimulus dimension, then combines SNRs among display modes, dimensions, and cognitive factors. In so doing, it can predict display format effectiveness. Examples of common metric assessment and validation are presented along with the derivation of the metric. Implications of the common metric in display design and evaluation are discussed.

Marshak. W. P., Winkler, R., Fiebig, C., Khakshour, A., & Stein, R. (1999)
Evaluating intelligent aiding of course of action decisions using the fox genetic algorithm in 2-D and 3-D interfaces
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 27–31

Intelligent aiding to improve decision processes and reduce support staff will become increasingly important in future Army tactical operations centers (TOCs). Federated Laboratory researchers have developed the Fox genetic algorithm (Fox-GA) decision aid to increase the number and quality of alternate courses of action (COAs) considered by the commander. Eleven Army officers at Fort Leavenworth, Kansas, used both traditional paper-based briefing and the Fox-GA COA generator to determine a COA in three different combat scenarios. Presentation of the Fox-GA COAs was made either within a two-dimensional (2-D) interface based on the ARL's CIP or within the National Center for Supercomputer Applications' BattleView three-dimensional visualization system. The findings indicate that Fox-GA significantly increased (by two to three times) the number of alternatives considered over the paper condition and that the 2-D visualization with Fox was both preferred and led to the

best performance. These results indicate that an improved GA-based COA generation system can significantly increase the number o f alternatives considered in the military planning process.

Martin-Emerson, R., & Kramer, A. F. (1997)
Offset transients modulate attentional capture by sudden onsets
Perception and Psychophysics, **59,** 739–751

> Recent research with visual search tasks has suggested that stimuli that appear as sudden onsets (new objects) have attentional priority over stimuli that are created by the removal of segments of premasks (non-onset stimuli). Attentional capture by sudden onsets occurs despite the fact that the appearance of these new objects predicts neither the identity nor the location of the target in the visual search task. In three experiments, we examined the extent to which attentional capture by sudden onsets could be modulated by offset transients used to create non-onset objects. To that end, we systematically manipulated the ratio of non-onset to onset stimuli in the display (display ratio) as well as the ratio of offset to onset segments between the stimulus types (stimulus ratio). Increases in either the stimulus ratio or the display ratio resulted in increases in the visual search slopes for the onset targets. These results suggest that the ability of sudden onsets (new objects) to capture attention is influenced by stimulus-driven factors, such as environmental change. Interestingly, the results also indicate that goal-directed or purposeful search for sudden-onset (new object) targets was relatively uninfluenced by the amount of change in the visual display. Therefore, it would appear that environmental change has differential effects on goal-directed and stimulus-driven search. These results are discussed in terms of their implications for our understanding of attentional capture.

Marzen, V., Stuppi, A., & Parent, J. (1997)
Display and control devices for advanced human computer interface
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 85–94

> Many advances have been made in display-and-control technology relating to the human-computer interface. This paper examines the devices and the techniques used to display and interact with computer-generated information. Direct view, projection, and body-worn displays will be evaluated for their ability to present information in traditional, immersive, and augmented environments. Similar comparisons will be made for control devices such as tactile, voice, and gesture.

McCarley, J. S., Kramer, A. F., & Peterson, M. S. (2001)
Object-based control of overt attention
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 47–52

> It is well established that control of covert attention is constrained by perceptual organization, so that attention spreads more easily within a

63

single object than between objects. We conducted two experiments to examine the role of perceptual organization in the control of overt attention shifts. Stimuli were pairs of adjacent elongated rectangles. Observers were asked to make a speeded judgment of the orientation of a target character appearing inside one rectangle, and a cue was provided before target onset to indicate the target's likely location. Gaze-contingent presentation of target and distracters was used to encourage eye movements. Eye movements during task performance evinced two forms of object-based effects. First, saccades following fixation of an invalidly cued item were more likely to be made within the cued rectangle than between rectangles. Second, saccades within the cued rectangle were preceded by shorter dwell times than saccades between rectangles. Data indicate that the control of overt attention is sensitive to the perceptual organization of a display.

McConkie, G. W., & Loschky, L. C. (1997)
Human performance with a gaze-linked multi-resolutional display
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 25–34

One method of reducing bandwidth requirements for displays is to present high-resolution information only at the location where the observer's gaze is directed. Two studies are reported that investigate the size to which the high-resolution "window" can shrink and the degree to which information outside this window can degrade without human performance being detrimentally unaffected.

McConkie, G. W., & Loschky, L. C. (2000)
Attending to objects in a complex display
Proceedings of the 4th Annual Federated Laboratory Symposium, 21–25

In the large virtual reality environments being developed for the military, personnel are faced with complex, dynamic displays containing many objects and regions, many of which reside outside the observers' field of view. Observers must form a mental representation of this space, remembering the relative positions of important objects, in order to be able to locate information quickly when needed. They then must monitor changes in this configuration in order to track the evolution of a battle. We are studying the perceptual processes involved in accomplishing these tasks.

McConkie, G. W., Loschky, L. C., Wolverton, G. S. (2000)
How well can binocular eyetracking indicate the depth plane on which attention is focused?
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 173

Eye tracking in a virtual environment is very useful in indicating what an observer is attending to at any given moment. However, in virtual three-dimensional environments, it is quite possible to have different objects

lying roughly in the same direction but at different depths from the observer. One method for obtaining information about the observer's depth plane of attention is through binocular eye movement recording. The angles of the two eyes with respect to each other change as attention is shifted between near and far objects. Attempts to use vergence in this way have not been very successful, and the question arises whether the failure is attributable to unreliability in people's eye positioning or poor accuracy of the eye-racking devices used by the researchers. Results of the present poster indicate that past failures are mainly attributable to eye tracker inaccuracy.

McConkie, G., & Rudmann, D. S. (1998)
Acquiring spatial knowledge from varying fields of view
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 49–53

> One effect of digitizing Army information is that commanders and their staffs often view a large battle space through a computer monitor that shows only part of the space at once. A study was conducted to examine the effect of field of view or viewport size on a person's ability to develop and use a mental representation of objects in a large terrain. Smaller viewports increase error in finding previously seen objects and remembering where they are located but do not affect simple memory for those objects.

McConkie, G. W., Zheng, X. S., & Schaeffer, B. (2001)
Effects of navigation control method on spatial updating in virtual environments
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 59–64

> We investigated the effects of different navigation control methods on human mental representations of virtual environments. Three control devices, joystick, wand and head tracker, in two different modes, absolute versus relative, were used to test hypotheses that absolute and more egocentric control devices would produce higher quality spatial mental representations than relative and non-egocentric devices. Results indicated an advantage for absolute mode devices in comparison to relative mode but no benefit for egocentric devices.

McCormick, E. R., Wickens, C. D., Banks, R. & Yeh, M. (1998)
Frame of reference effects on scientific visualization subtasks
Human Factors, 40, 443–451

> Performance measures for three frames of reference (full egocentric, full exocentric, and tethered) were contrasted across four different scientific visualization subtasks: search, travel, local judgment support, and global judgment support. Participants were instructed to locate and follow a designated path through 15 simple virtual environments and answer simple questions about that environment. Each participant completed

five trials in all three frame-of-reference conditions. The results revealed that frames of reference that use egocentric or tethered viewpoints support better travel performance, especially when participants were nearing the target. However, the exocentric frame of reference supported better performance in the search subtasks and in the local and global judgment subtasks. Actual or potential applications of this research include proper uses of virtual reality to support certain scientific visualization subtasks.

McGee, J. H., Chen, S. L., Sundareswaran, V. S., Vassiliou, M. S., & Marshak, W. P. (2001)
Comparing pointing, speech, and combined point-and-speak control input
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 11–15

> Research at the Rockwell Scientific Company (RSC) has suggested the superiority of combined pointing and speaking as a software control input in certain conditions. To quantify the relative effectiveness of pointing, speech, and combined point-and-speak control input, RSC developed new experiment software built on previous RSC research. A pilot study of human subjects conducted by Sytronics used this software to yield quantifiable comparison data. The results of these trials provide an initial step in the comparison of input modalities in controlled conditions.

Mehrotra, S., Rui, Y., Chakrabarti, K., Ortega, M., & Huang, T. S. (1997, September)
Multimedia analysis and retrieval system
Paper presented at the meeting of the 3rd International Workshop on Information Retrieval Systems, Como, Italy

> No abstract

Mehrotra, S., Rui, Y., Ortega-Binderberger, M., & Huang, T. S. (1997)
Supporting content-based queries over images in MARS
Proceedings of the IEEE International Conference on Multimedia Computing and Systems '97, 632–633

> While advances in technology allow us to generate, transmit, and store large quantities of digital images, video, and audio, research in the indexing and retrieval of multimedia information is still in its infancy. To address the challenges in building an effective multimedia database system, we have built the multimedia analysis and retrieval system (MARS) prototype. This paper summarizes the retrieval subsystem of MARS and how it supports content-based queries over image features. Content-based retrieval techniques have been extensively studied for textual documents in the area of automatic information retrieval. Our objective in MARS is to exploit these existing techniques for content-based retrieval over images and multimedia databases.

Mengshoel, O. J. (1997)
Belief network inference in dynamic environments

Proceedings of the 14th national conference on Artificial Intelligence, 813

No abstract

Mengshoel, O. J. (1999)
Evolutionary computation in Bayesian networks

in J. R. Koza (Ed.), *Late Breaking Papers at the Third Annual Genetic Programming Conference on System Sciences* (p 159). Madison, WI: Omni Press

Genetic algorithms (GAs) are stochastic algorithms for search, optimization, and machine learning. In this research, the focus is on using a Bayesian network (BN) as the GA fitness function. More formally, a Bayesian network is a tuple (V, W, P_r), in which (V, W) is a directed acyclic graph with nodes $V = \{V_1, ...V_n\}$ and edges $W = \{W_1...W_m\}$; P_r is a set of conditional probability distribution tables. The nodes correspond to random variables and the edges to conditional dependencies between these random variables. For each node $V_i \in V$, there is one conditional probability table that defines a conditional probability distribution over V_i in terms of its parents $P_a(V_i)$: $P_r(V_i \mid P_a(V_i)) \in P_r$.

Mengshoel, O. J., & Goldberg, D.E. (1999, July)
Probabilistic crowding: Deterministic crowding with probabilistic replacement

paper presented at the 1999 *Genetic and Evolutionary Computation Conference*, Orlando, FL

This paper presents a novel niching algorithm: probabilistic crowding. Like its predecessor (deterministic crowding), probabilistic crowding is fast and simple, requiring no parameters beyond those of the classical genetic algorithm. In probabilistic crowding, sub-populations are maintained reliably, and we analyze and predict how this maintenance takes place. This paper also identifies probabilistic crowding as a member of a family of algorithms that we call integrated tournament algorithms. Integrated tournament algorithms also include deterministic crowding, restricted tournament selection, elitist recombination, parallel recombinative simulated annealing, the Metropolis algorithm, and simulated annealing.

Mengshoel, O. J., Goldberg, D. E., & Wilkins, D. C. (1998)
Deceptive and other functions of unitation as Bayesian networks

In J. R. Koza (Ed.), *Genetic Programming*, (pp 559–566). San Francisco, CA: Morgan Kaufmann

In trying to understand which fitness functions are hard and which are easy for genetic algorithms (GAs) to optimize, researchers have considered deceptive and other functions of unitation. This paper focuses on GA fitness functions represented as Bayesian networks. We investigate onemax, trap, and hill functions of unitation when they are converted

into Bayesian networks. This paper shows, among other things, that Bayesian networks can be deceptive.

Mengshoel, O. J., Roth, D., & Wilkins, D. C. (2000)
Hard and easy Bayesian networks for computing the most probable explanation
Tech. Rep. No. UIUC DCS-R-2000-2147, University of Illinois at Urbana-Champaign, Computer Science Department

This paper introduces an experimental paradigm for systematically generating increasingly hard instances for Bayesian network inference. The approach allows us to control the level of difficulty of the Bayesian network inference problem, providing benchmark Bayesian networks for more systematic experimentation. We investigate two families of synthetic Bayesian networks, in which we study a few structural and distributional parameters and show how changing them (while maintaining network size) can change the hardness of the problem from a very simple inference problem to one that existing algorithms cannot handle. Among the parameters we study are the ratio of the number of root nodes to the number on non-root nodes in the network, the irregularity of the graph, and the distributional nature of the conditional probability tables. The difficulty of the networks is investigated experimentally via one of the most successful commercial inference algorithms, *Hugin*, along with a stochastic local search algorithm that we have developed: *stochastic greedy search*. While both algorithms degrade as the difficulty of the problem increases, we show that they vary significantly along some of the dimensions and that, surprisingly, the performance of the stochastic search algorithm degrades more gracefully in many cases.

Mengshoel, O. J., & Wilkins, D. C. (1996)
Recognition and critiquing of erroneous agent actions
Proceedings of the American Association for Artificial Intelligence, Workshop on Agent Modeling, 61–68

No abstract

Mengshoel, O. J., & Wilkins, D. C. (1996)
Toward an approach to exploiting domain structure for planning
Presented at the AAAI-96 Workshop on Structural Issues in Planning and Temporal Reasoning

No abstract

Mengshoel, O. J., & Wilkins, D. C. (1997)
Abstraction and aggregation in belief networks
Proceedings of the Workshop for Abstraction, Decisions, and Uncertainty at the 14th National Conference on American Association for Artificial Intelligence, 53–58

Abstraction and aggregation are useful for increasing speed of inference in and easing knowledge acquisition of belief networks. This paper

presents previous research on belief network abstraction and aggregation, discusses its limitations, and outlines directions for future research.

Mengshoel, O. J., & Wilkins, D. C. (1997)
Visualizing uncertainty in battlefield reasoning using belief networks

Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. P), 15–22

> Battlefield reasoning is a complex reasoning task in which uncertain and incomplete knowledge is crucial, particularly regarding enemy activity. One approach to uncertainty reasoning in artificial intelligence is belief networks. Belief networks have a graph structure that also facilitates visualization. For these reasons, we suggest belief networks as a central knowledge representation formalism for the battlefield reasoning task. Belief networks are useful for doing data fusion in the presence of uncertainty. Data fusion merges information from diverse information sources (or sensors) with varying reliability or probability of failure. In the battlefield reasoning task, the information sources would be human and automated intelligence assets. While belief networks are fundamentally well suited to the battlefield reasoning task, more research is needed about temporal reasoning via dynamic belief networks. This paper presents a basis dynamic belief network model and proposes to extend it to an event-based approach to dynamic belief network representation and reasoning. This preliminary model is based on an analysis of the battlefield reasoning task.

Mengshoel, O. J. & Wilkins, D.C. (1998, March)
Abstraction for belief revision: Using a genetic algorithm to compute the most probable explanation

paper presented at the AAAI Spring Symposium Series, Stanford University, Menlo Park, CA

> A belief network can create a compelling model of an agent's uncertain environment. Exact belief network inference, including computing the most probable explanation, can be computationally difficult. Therefore, it is interesting to perform inferences on an approximate belief network rather than on the original belief network. This paper focuses on approximation in the form of abstraction. In particular, we show how a genetic algorithm (GA) can search for the most probable explanation in an abstracted belief network. Because belief network approximation can be treated as noise from the point of view of a GA, this topic is related to research on noisy fitness functions used for GAs.

Mengshoel, O. J., & Wilkins, D. C. (1998)
Genetic algorithms for belief network inference: The role of scaling and niching

in V. W. Porto, N. Saravanan, D. Waagen, & A. E. Eiben (Eds.), *Proceedings of the 7th International Conference on Evolutionary Programming* (pp 547–556). Berlin, Germany: Springer-Verlag

> Belief networks encode joint probability distribution functions and can be used as fitness functions in genetic algorithms (GAs). Individuals in the GA's population then represent instantiations or explanations in the belief network. Computing the most probable explanations (belief revision) is thus cast as a GA search in the joint probability distribution space. At any time, the best fit individual in the GA population is an estimate of the most probable explanation. This paper argues that joint probability distribution functions represented by belief networks typically are multimodal and highly variable. Thus, the GA techniques known as sharing and scaling should be helpful. It is shown empirically that this is indeed the case, particularly that niching combined with scaling significantly improves the quality of a GA's estimate of the most probable explanations. A novel scaling approach, root scaling, is also introduced.

Mengshoel, O. J., Wilkins, D. C., & Uckun, S. (1998)
Filtering and visualizing uncertain battlefield data using Bayesian networks

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 74–78

> Filtering, interpreting, and visualizing massive amounts of uncertain data are a core challenge in battlefield reasoning. Another challenge concerns the uncertain and incomplete knowledge about enemy and even friendly forces. This paper presents a Bayesian network approach as a way to deal with these challenges. We present Bayesian networks and describe how they can be used for battlefield reasoning, particularly intelligence analysis. We emphasize how Bayesian networks can be used for intelligent information processing in the form of filtering, fusion, and selection of information.

Merlo, J. L., Wickens, C. D., & Yeh, M. (1999)
Effect of reliability on cue effectiveness and display signaling

Tech. Rep. No. ARL-99-4/FED-LAB-99-3, Urbana-Champaign: University of Illinois, Aviation Research Lab, Institute of Aviation

> The effects of automation failure on trust and of visual cuing precision on attention were investigated in a target detection task. Twenty military subjects searched a simulated mountainous terrain for military-relevant targets while performing a secondary monitoring task on either a hand-held display (HHD) or a helmet-mounted display. Both displays had target cuing present for half the trials, with the precision of the target cues

varied across blocks. Cued trials were either precise (a cuing reticle always circumscribed a target) or imprecise (the target was outside the reticle by 22.5° or 45°). Imprecise cuing simulated degraded sensor resolution. Cue precision and imprecision were conveyed to subjects by solid or dashed lines, respectively. A high-priority target was presented twice each block, once with a precisely cued target and once with an imprecisely cued target. Target cuing induced an attention cost (as revealed by the low detection rate of high-priority uncued targets), when a cue occurred simultaneously with a low-priority target. During the last experimental block, the automated target cuing failed on some trials, resulting in attention and trust costs, with subjects initially showing signs of over-trust of the cuing information and then on subsequent trails tending to under-trust the cuing information, with trust seemingly restored after a few reliable trials. Failures in automation also seemed to mediate the effects of attention costs, as the detection rate of the higher priority but uncued target increased.

Merlo, J. L., Wickens, C. D., & Yeh, M. (2000)
Effect of reliability on cue effectiveness and display signaling
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 27–31

Twenty Army personnel detected, identified, and reported the azimuth of targets in scenes projected on a three-walled video environment. Target cuing, presented on either a hand-held display or a helmet-mounted display, occurred for half of the trials, with the precision of the cues varying across blocks of trials. In most trials, a single target was presented, but in 10% of the trials, a second target was also presented. In cued trials, subjects tended to miss the second target, focusing their attention on the area around the cue. During the last experimental block, target cuing failed on some trials, after which, subjects exhibiting over-trust in automated cuing prolonged their search of an area for a target. However, failures in automation seemed to expand subjects' search area for a target in subsequent trials with restored cue reliability, in that subjects' detection of the second target increased.

Molineros, J., Raghavan, V., & Sharma, R. (1999)
AREAS: Augmented reality for evaluating assembly sequences
in Behringer, R., Klinker, G., & Mizell, D. W. (Eds.). *Augmented reality: Placing artificial objects in real scenes* (pp 91-99). Natick, MA: A. K. Peters

Augmented reality provides a powerful and intuitive interface that can enhance the user's understanding of a scene. We consider the problem of scene augmentation in the context of the assembly of a mechanical object. Concepts from robot assembly planning are used to develop a systematic framework for presenting augmentation stimuli for this assembly domain. We then describe an interactive evaluation tool called AREAS, which uses augmentation schemes for visualizing and evaluating assembly sequences. The system also guides the user step by step through

an assembly sequence. Computer vision, together with a system of markers, provides the sensing mechanism necessary to interpret the assembly scene.

Mountjoy, D. N., Chi, C. J., Ntuen, C. A., & Yarbrough, P. L. (1997)
Associative configural display of dynamic tactical information
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. P), 53–56

This paper describes the concept of an associative configural display (ACD) and how it might be applied in tactical decision making on the battlefield. ACD provides dynamic summary information regarding unit effectiveness during the course of a mission and compares actual effectiveness with the commander's original plan. The configuration of the display would allow a commander to make fast and accurate decisions regarding unit status and could be used to track multiple units concurrently.

Mountjoy, D. N., & Marshak, W. (1999)
Impact of non-linear mapping on mileage estimation
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 97–101

Nonlinear mapping is a display technique that can be applied to situation maps to maintain detail in the commander's area of interest while displaying more peripheral land area to convey contextual information. A series of studies has been undertaken to explore the perceptual advantages and limitations of this technique in an effort to produce a more efficient tactical mapping system. The first of this series (the effect on mileage estimation) is discussed here, along with directions of future research.

Mountjoy, D. N., Marshak, W. P. Converse, S. A., & Ntuen, C. A. (2000)
Perception and performance effects of nonlinear map representations
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 175

Small electronic displays typically found in aircraft cockpits and other vehicular applications (e.g., a command and control vehicle) must trade area coverage and detail because of low pixel density and physical space constraints. One approach for increasing area coverage is to exploit the "elasticity" of electronic displays by the use of nonlinear scaling. For example, the undistorted area of interest on a map can be placed in the center of an electronic display, and the surrounding area can be squeezed to fit on the periphery of the display. By applying this technique, one can maintain the display of contextual information necessary for preserving a strong sense of situation awareness. Simultaneously, nonlinear maps should lessen the required number of interface interactions since detail and context can be provided on the same map surface. However, possible adverse effects on navigation tasks may lessen the overall benefit of these

nonlinear representations. Three experiments have been designed to examine the effects of nonlinear mapping on mileage and heading estimation and to examine the proposed benefits of battlefield monitoring performance. Results of this research are intended to help guide the development of efficient, cost-effective, small-screen tactical displays.

Mountjoy, D. N., Marshak, M. P., & Ntuen, C. A. (2001)
Performance evaluation of a perception-based non-linear map display
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 23–28

> This paper compares human performance when two types of electronic maps are used: a nonlinear tactical map and a linear map configured with a standard pan and zoom interface. No performance differences were found between the two maps in navigational accuracy or response time. Subjective workload was also unaffected by map type. However, the expected gain in detecting randomly occurring events (through display of additional context) was not evident. Suggestions about the cause of this "non-finding" are offered.

Munetomo, M., & Goldberg, D. E. (1999)
Identifying linkage groups by nonlinearity/non-monotonicity detection
Proceedings of the 1999 Genetic and Evolutionary Computation Conference, 433–440

> This paper presents and discusses direct linkage identification procedures based on nonlinearity/non-monotonicity detection. The algorithm we propose checks arbitrary nonlinearity/non-monotonicity of fitness change by perturbations in a pair of loci to detect their linkage. We first discuss the condition of the LINC (linkage identification by a nonlinearity check) procedure and its allowable nonlinearity. Then we propose another condition of the LIND (linkage identification by non-monotonicity detection) and prove its equality to the LINC with allowable nonlinearity (LINC-AN). The procedures can identify linkage groups for problems with (at most) order-k difficulty by checking $O(2^k)$ strings; the computational cost for each string is $O(l^2)$, in which l is the string length.

Naphade, M. R., & Huang, T. S. (2001)
A probabilistic framework for recognizing audio-visual semantics
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 79–84

> Video content is an important medium in battlefield communication. Access to this content, however, is far from efficient. The most natural and user-friendly access mechanism to video is semantic keywords. These keywords should ideally represent various semantic concepts such as objects, sites, and events. Automatic annotation of video by these keywords is very difficult. For this, we need models that represent these keywords in multimodal feature spaces. For many interesting and useful concepts, it may be possible for annotation software to learn such models

from training data. This paper proposes a probabilistic framework for semantic video indexing. The components of the framework are multijects and multinets. Multijects are probabilistic multimedia objects representing semantic features or concepts. A multinet is a probabilistic network of multijects, which accounts for the interaction between concepts. Using the framework, we show how semantic objects such as "human presence" can be modeled. Results indicate the importance of multi-modality in detecting such concepts.

Ntuen, C. A. (1999)
An ecological model of situation awareness: What does it mean to battlefield awareness?

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 153

Most studies of situation awareness (SA), especially those designed for decision aiding, rely on the theories and models of cognition and perception. Theories developed by Endsley and by Pew conceptualize SA as the interaction of product and process. *Product* refers to the state of our knowledge about the environment, and *process* refers to the perceptual and cognitive activities that update our knowledge. The author discusses how these ideas pertain to designing decision-aiding software applications.

Ntuen, C. A., Chi, C.-J., McBride, M. E., & Park, E. H. (1998)
Decision support display modeling for digital battlefield

Proceedings Fourth Annual Symposium on Human Interaction with Complex Systems, 155–159

A decision support display (DSD) was developed as a cognitive aiding tool to support the decision maker in an unstructured, dynamic, uncertain, and information-intensive environment. Battlefield information is modeled as a context-dependent and action-oriented object that adapts to a defined system goal or mission statement. The DSD philosophy is applied to a graphical display of alternate courses of action designed to amplify the decision maker's knowledge and experience levels.

Ntuen, C. A., & Deng, F. (2000)
Evaluating multimodal interface performance with human operator control models

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 65–69

An integrated approach to modeling human performance in a closed loop, multimodal information processing system is developed and validated. The technique blends human response theory and modern control theory to analyze potential human response performance during task control in a sensory-information-processing environment. The model dynamics include visual displays and tactile and auditory information

74

presentation. The human response is considered a function of the system state variables. We have developed a simulation model for use in constructive experiments for determining the effects of multimodal information processing on human performance. The control simulation is generic and thus useful as a common metric for evaluating system performance.

Ntuen, C. A., Mountjoy, D. N., Barnes, M. J., & Yarborough, L. P. (1997)
Representation of the commander's heuristic knowledge in a decision support display
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 1), 41–56

> In this paper, we discuss a framework for representing the commander's decision heuristics in a display environment. We use cognitive task analysis to assess the commander's heuristics at various levels of task abstraction. The knowledge-representation model is conceived to enhance the decision support display being developed for tactical command and control at the brigade level and below.

Ntuen C. A., Park E. H., Chi C., Yarborough L. P., & Mountjoy D. N. (1999)
Effect of information presentation mode on condition monitoring of battlefield events
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 155

> The goal of this study was to determine the most effective method of presenting critical battle information to the commanders to ensure the rapid detection of potentially disastrous conditions. Electronic map displays containing unit symbols and course of action arrows, which were drawn with bands across them, served as stimuli. Four methods of presentation were tested: color band changes; color band changes and flashing unit symbols; color band changes and an auditory alarm; color band changes, flashing unit symbols, and an auditory alarm. Results indicated that performance was faster in the second and fourth conditions than in the first and third.

Ntuen, C. A., Park, E. H., Chi, C., Yarborough, L. P., & Mountjoy, D, N. (1999)
Human performance with decision support display
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 157

> A laboratory experiment was conducted to evaluate human decision-making performance when the Alternative Courses of Action Display (ACAD) software was used. Two general tasks associated with information visualization were tested: extraction of information and decision tasks. Information extraction depended on the user's cognitive factors, which are affected by the realism of display cues (i.e., how closely the ACAD represents the physical objects), correlation between information on the display with mental models (a measure of how closely

the ACAD representation of physical objects matches what the user already knows about the objects), and reminders (a measure of how well display cues improve recall of information from memory). The decision and execution tasks studied were feature detection and recognition of battle events. Feature detection concerns the user's ability to detect changes in the object states, based on a display scenario. Recognition of battle refers to the ability of the user to recognize salient decision variables in a display. In order to determine the strength of each critical element, a laboratory experiment was conducted to determine the correlation between the three levels of information extraction criteria (cognitive fit tasks) and the individual components of decision tasks.

Ntuen, C. A. Park, E. H., Eastman, S., Mountjoy, D., & Yarbrough, L. P. (2000)
ACAD: A decision support display for commander's visualization of alternative courses of action during battle planning
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 101–105

This paper describes a decision support display, called alternative courses of action display (ACAD), designed to support the commander's battle-planning and course of action (COA) analysis. ACAD is a planning and experimental decision-making tool that contains information about the battle situation, the resources available, and the enemy's situation. Because the military commander must compare friendly COAs with enemy COAs, a common performance measure of effectiveness used by ACAD is the relative force ratio, a relative measure of friendly force strength against the enemy's force strength. We also show some results of pilot usability analysis of ACAD.

Ntuen, C. A., Park, E. H., Evans, M., Borhauer, R., Hocking, D., Leininger, J., & Harder, R. (1998)
Human factors issues in collaborative planning
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 120–124

Collaborative planning involves the use of multiple (intelligent) agents in a problem-solving team in a domain-specific environment. Human-human collaborative planning predominates in the military decision-making process. However, with the recent progress in human-computer interaction, computer-supported cooperative work, and group decision support systems, military decision making will be more automated, thereby requiring a mix of humans interacting on a more cognitive level with intelligent software agents. This brings some theoretical issues related to representation of human factors elements into collaborative planning.

Nwankwo, H., Deol, D., Aikens, S., Goodwin-Johansson, S., & Marshak, W. (2000)
Experiments to determine efficacy of a tactile interface coding strategy

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 177

> This paper reports the results of experiments conducted to determine the efficacy of coding schemes for efficient and effective transmission of situation data via a tactile interface. A tactile interface is useful for transmitting information to individual soldiers who must use their visual and auditory systems to monitor the surrounding environment. Of interest to the researchers was the extent to which the tactile modality can be engaged reliably, efficiently, and effectively as a communications tool. The study design, considerations, and assumptions, as well as the resulting data and inferences drawn, are presented.

Nwankwo, H. E., Goodwin-Johansson, S. H., & Mancusi, J. E. (1997)
Tactile interface: Cognitive, psychophysical, and physiological issues

Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 95–105

> In this paper, we discuss the cognitive, psychophysical, and physiological issues that must be considered in the development of specifications for the design of a tactile interface. In general, cognitive issues pertain to how tactile information should be presented so that it is easily understood by a user. Physiological and psychophysical factors are related and pertain to the optimum location of the tactile device on the body and the intensity necessary for the stimulation to be detected.

Nwankwo, H. E., Urquhart, R., Goodwin-Johansson, S., & Mancusi J. (1999)
Tactile communication interface design: Efficacy of euphemistic terms as interface location cues

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 159

> To apply tactile interface communication for the purpose of increasing human information processing, we must address the issue of how the interface device should be designed to ensure meaningful information transfer. In this paper, we examine the relationship between a set of military communications (e.g., danger area, stop) and associated body locations (e.g., armpit) and gestures (e.g., "cut throat"). Eighty subjects indicated on a questionnaire how intuitive the relationship between a military communication and body location was. For example, "danger area" was strongly related to "armpit" and "stop" to a "cut throat" gesture. The body locations identified could become interface locations for receiving tactile messages. Experiments are under way to validate findings gleaned from subjects' questionnaire responses.

Ortega, M., Chakrabarti, K., Porkaew, K. & Mehrotra, S. (1998, June)
Cross media validation in a multimedia retrieval system

Paper presented at the 3rd ACM Conference on Digital Libraries, Digital Library Metrics Workshop, Pittsburgh, PA

> The increasing size of document databases has prompted a change from manual indexing and querying to automated methods. This switch necessitated a performance metric for the automated systems; however, performance measurement of automated systems was and still is performed manually. Ever-increasing collection size makes manual evaluation progressively more difficult, and this difficulty is compounded by the addition of multimedia. In this paper, we describe an automated method for measuring the retrieval performance of a new arbitrary retrieval algorithm suited to a particular media type.

Ortega, M., Rui, Y., Chakrabarti, K., Mehrotra, S., & Huang, T. S. (1998)
Supporting similarity queries in MARS

Proceedings of the Fifth ACM International Multimedia Conference, 403–413

> To address the emerging needs of applications that require access to and retrieval of multimedia objects, we are developing the multimedia analysis and retrieval system (MARS). In this paper, we concentrate on the retrieval subsystem of MARS and its support for content-based queries over databases containing images. Content-based retrieval techniques have been studied extensively as a means of automatic information retrieval of documents containing textual material. This paper describes how these techniques can be adapted for ranked retrieval over image databases. We focus on MARS's Boolean retrieval model and describe the results of our experiments demonstrating the effectiveness of the model for image retrieval.

Ortega, M., Rui, Y., Chakrabarti, K., Porkaew, K., Mehrotra, S., Huang, T. S. (1999)
Supporting ranked Boolean similarity queries in MARS

IEEE Transactions on Knowledge and Data Engineering, **10,** 905–925

> To address the emerging needs of applications that require access to and retrieval of multimedia objects, we are developing the multimedia analysis and retrieval system (MARS). In this paper, we concentrate on the retrieval subsystem of MARS and its support for content-based queries over image databases. Content-based retrieval techniques have been extensively studied for textual documents in the area of automatic information retrieval. This paper describes how these techniques can be adapted for ranked retrieval over image databases. Specifically, we discuss the ranking and retrieval algorithms developed in MARS, based on the Boolean retrieval model and describe the results of our experiments, which demonstrate the effectiveness of the developed model for image retrieval.

Ortega-Binderberger, M., Mehrotra, S., Chakrabarti, K., & Porkaew, K. (2000, January)
WebMARS: A multimedia search engine

Proceeding of the International Society of Optical Engineering, *314–321. (Also a Technical Report, No. TR-MARS-2000-01, University of California-Irvine)*

> Describes WebMARS, a search engine that uses textual and visual information for hypertext markup language (HTML) document retrieval. Textual information can take the form of words or citations. Visual information can be simple (color, texture, or image patterns) or more complex (organization of color or patterns). The ability to refine a query, based on the results of a search, is implemented in the system.

Oswald, S. P., Ramchandran, K., & Huang, T. S. (1997, September)
Efficient terrain data representation for 3D rendering using the generalized BFOS algorithm

Proceedings of the International Conference on Image Processing, *1, 448–451*

> Digital terrain data have widespread applications in areas such as military virtual battlefields, geographic information systems (GIS), and flight simulator video games. The combination of the abundance of terrain data with the limited rendering capabilities of computer graphics equipment creates the necessity for algorithms that generate efficient representations of the data for rendering. This paper presents such an algorithm. Terrain data are represented by a binary tree, and the generalized Breiman, Friedman, Olshen, and Stone (BFOS) algorithm, a well-known optimal tree-pruning method for regression and quantization trees, is used to optimally prune the tree, resulting in a far more efficient representation of terrain data than has previously been attained.

Pavlovic, V. I., Berry, G. A., Huang, T. S., Devi, L., Sethi, Y., & Sharma, R. (1998)
Speech/gesture integration for display control

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, *79–84*

> Although computer technology has dramatically changed in the last 20 years, human-computer interfaces have largely remained the same. The keyboard has been the most prevalent device, but with the advent of graphical operating systems, the mouse has been added. To create a more natural and human-centric computer interface, we propose using input modalities that are employed in daily human communications. By replacing the keyboard and mouse with a gesture and speech recognition system, we can develop more natural controls for numerous application. In this paper, we explore the use of speech and gesture modalities in a display control application.

Pavlovic, V., & Huang, T. S. (1999)
Multimodal prediction and classification of hand gestures and speech

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 161

> The authors propose a novel framework for multimodal feature
> prediction and classification based on multimodal hidden Markov
> models (MHMMs). Previous approaches employed loosely coupled
> unimodal techniques in which feature estimation, prediction, and lower
> level classification are performed independently within each of the
> modality domains. MHMMs model the redundancy among co-occurring
> modalities such as speech, hand gestures, lip motion, etc. In this report,
> the test bed application was a joint audio-visual interpretation of speech
> and unencumbered hand gestures for interaction with virtual
> environments. The setup allowed a user to interact with a three-
> dimensional virtual environment via hand gestures (such as pointing and
> simple symbolic motions) and spoken commands. Bimodal HMMs were
> employed to model the influence of speech on gestural actions. MHMM
> parameter learning was performed on a set of 39 bimodal commands. The
> test set was a different sequence of 31 commands performed by the same
> user. Two experiments compared the performance of bimodal with
> unimodal models on the test data. In the normal visual noise
> environment, recognition performance of bimodal HMMs significantly
> exceeds the performance of unimodal HMMs (62% versus 35%). High
> visual noise reduced the recognition performance of both models.
> However, bimodal HMMs retained a relatively significant recognition
> ratio of 52%, while the unimodal approach failed almost completely
> (10%). Results of the test indicated that the bimodal HMMs significantly
> improved the recognition performance in two different gestural speech
> classification tasks. Future work is aimed at further examination of the
> robustness of classification as well as the on-line implementation of the
> algorithms.

Pavlovic, V. I., Sharma, R., & Huang, T. S. (1996)
Gestural interface to a visual computing environment for molecular
biologists

Proceeding of the Second International Conference on Automatic Face and Gesture
Recognition, 30–35

> In recent years, there has been tremendous progress in three-dimensional
> (3-D) immersive displays and virtual reality (VR) technologies. Scientific
> visualization of data is one of many applications that has benefited from
> this progress. To fully exploit the potential of these applications in the
> new environment, there is a need for natural interfaces that allow the
> manipulation of such displays without burdensome attachments. This
> paper describes the use of visual hand gesture analysis enhanced with
> speech recognition for developing a bimodal gesture-speech interface for
> controlling a 3-D display. The interface augments an existing application,

VMD, which is a VR visual computing environment for molecular biologists. Hand gestures and a set of speech commands are used for manipulating the 3-D graphical display. We concentrate on the visual gesture analysis techniques used in developing this interface. The dual modality of gesture and speech greatly aids the interaction capability.

Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997)
Visual interpretation of hand gestures for human-computer interaction: A review
IEEE Transactions on Pattern Analysis and Machine Intelligence, **19,** 677–695

The use of hand gestures is an attractive alternative to cumbersome interface devices for human-computer interaction (HCI). In particular, visual interpretation of hand gestures can help provide the ease and naturalness desired for HCI. This has motivated active research in computer vision-based analysis and interpretation of hand gestures. In our review of the literature about visual interpretation of hand gestures in the context of its role in HCI, we organize our discussion according to the method used for modeling, analyzing, and recognizing gestures. Important differences in approaches to gesture interpretation arise depending on whether a 3-D model or an image appearance model of the human hand is used. Three-dimensional hand models allow more elaborate modeling of hand gestures but also lead to computational hurdles that have not been overcome, given the real-time requirements of HCI. Appearance-based models lead to computationally efficient "purposive" approaches that work well in constrained situations but seem to lack the generality desirable for HCI. We discuss implemented gestural systems as well as other potential applications of vision-based gesture recognition. Although the current progress is encouraging, further theoretical as well as computational advances are needed before gestures can be widely used for HCI. We also discuss directions of future research in gesture recognition, including its integration with other natural modes of human-computer interaction.

Pelikan, M., Goldberg, D. E., & Cantú-Paz, E. (1999)
BOA: The Bayesian optimization algorithm
Tech. Rep. No. 99003, Urbana-Champaign: University of Illinois, Illinois Genetic Algorithms Laboratory

We propose an algorithm that uses an estimation of the joint distribution of promising solutions to generate new candidate solutions. The proposed algorithm, based on the concept of genetic algorithms, is called the Bayesian optimization algorithm (BOA). To estimate the distribution of promising solutions, the algorithm exploits techniques for modeling multivariate data by Bayesian networks. The proposed algorithm identifies, reproduces, and mixes building blocks up to a specified order. It is independent of the ordering of the variables in the strings representing the solutions. Prior information about the problem can be incorporated into the algorithm, but it is not essential. Preliminary experiments show that as the problem size grows, the BOA outperforms

the simple genetic algorithm, even in decomposable functions with tight building blocks.

Peterson, M. S., & Kramer, A. F. (2001)
Guidance of the eyes by contextual information and abrupt onsets

Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 41–45

> Contextual cuing is a memory-based phenomenon in which previously encountered global pattern information in an environment can automatically guide attention to the location of a target, leading to rapid and accurate responses. Abrupt visual onsets have been shown to automatically capture attention and the eyes in situations that require eye movements. In real and virtual environments, memory-based and stimulus-driven guidance often compete to drive attention. In a series of experiments, we find that although contextual information can partially override capture by abrupt onsets, contextual cuing is a weak phenomenon that occurred only in some trials and at times not until later during the search process.

Peterson, M. S., Kramer, A. F., Irwin, D. W., Hahn, S. (2000)
Modulation of oculomotor capture by abrupt onsets during free viewing

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 131–134

> Abrupt visual onsets have been shown to automatically capture attention in situations that do not require eye movements. However, in real and virtual environments, eye movements are typically employed when one is scanning for information. In a series of experiments, we find that onset relevance and the degree of saccade planning can modulate the probability that visual onsets will capture attention in situations that require eye movements.

Poddar, I., Sethi, Y., Ozyildiz, E., & Sharma, R. (1998)
Toward natural gesture/speech HCI: A case study of weather narration

Proceedings of the Workshop on Perceptual User Interfaces (PUI'98), 1–6

> For human-computer interaction to be more natural, computers must be able to recognize continuous natural gestures and speech. To this end, previous researchers, using hidden Markov models (HMMs), have reported high recognition rates for gesture recognition; however, these gestures were defined precisely and were bound with syntactical and grammatical constraints. Natural gestures neither string together in syntactical bindings nor are amenable to strict classification. By recording the hand gestures and speech of a reporter standing before a weather map, we have studied the interaction between speech and gesture in the context of a display. We have implemented a continuous HMM-based gesture-recognition framework. To understand the interaction between gesture and speech, we conducted a co-occurrence analysis of different gestures with some spoken keywords. We also demonstrated the

possibility of improving continuous gesture recognition results, based on the co-occurrence analysis. Fast feature extraction and tracking are accomplished by the use of predictive Kalman filtering on a color-segmented stream of video images. The results in the weather domain should be a step toward a natural gesture-and-speech computer interface.

Poddar, I., & Sharma, R. (1999, November)
Continuous recognition of natural hand gestures for human computer interaction

Paper presented at the 12th Annual ACM Symposium on User Interface Software and Technology (UIST '99), Asheville, NC

> The use of hand gestures is an attractive alternative to cumbersome interface devices for human-computer interaction (HCI), particularly within a multimodal system, such as a speech and gesture interface. In particular, visual interpretation of hand gestures can help achieve the ease and naturalness desired for HCI. To exploit this potential, we need to develop recognition techniques that can handle continuous natural gesture input. Natural gestures are usually embedded in speech with no fixed, predefined meanings, and they do not string together in any syntactic bindings. In this paper, we propose techniques for the recognition of natural gestures that occur in the context of controlling and interacting with spatial maps through speech and gesture. We first present a study of a "parallel" domain using data from the weather narration in broadcast TV. This gives us a way to bootstrap the development of a gesture/speech system for interacting naturally with a graphical display of a spatial map.

Porkaew, K., Chakrabarti, K., & Mehrotra, S. (1999)
Query refinement for multimedia similarity retrieval in MARS
Proceeding of the 7th ACM International Multimedia Conference, 235–238

> A new method for refining queries in the multimedia analysis and retrieval system (MARS) was compared with a method already incorporated in MARS. The researchers posit a two-step process for multimedia searches. Users create initial queries by providing examples of objects similar to those they wish to retrieve; then, in a step called "relevance feedback," they modify their queries by indicating which of the returned objects is most like the objects they seek. An object is represented as a collection of features, which in turn are represented by vectors in an object space. A query is represented as the sum of several object spaces. During the relevance feedback step, a clustering technique called *query expansion* is used to modify a query by identifying a set of objects to be added to the query representation. Experimental results show that query expansion significantly outperforms an older query modification technique in MARS (query point movement), both in terms of retrieval effectiveness and execution costs.

Porkaew, K., Mehrotra, S., & Ortega, M. (1999)
Query reformulation for content based multimedia retrieval in MARS

*IEEE International Conference on Multimedia Computing and Systems, **2**, 747–751*

> Unlike traditional database management systems, content-based multimedia retrieval databases make it difficult for a user to ask for information in a direct, precise query. A typical multimedia interface allows a query to be based on examples of objects similar to the ones users wish to retrieve. Such an interface, however, requires mechanisms for the system to learn the query representation from the examples. In this paper, we describe the query refinement framework implemented in the multimedia analysis and retrieval system for learning query representations using relevance feedback. The proposed framework uses a query expansion approach to modifying the query representation, in which relevant objects are added to the query. Furthermore, query re-weighting techniques are used to adjust similarity functions.

Porkaew, K., Mehrotra, S., Ortega, M., & Chakrabarti, K. (1999)
Similarity search using multiple examples in MARS

*Lecture Notes in Computer Science, **1614**, 68–75*

> Unlike traditional database management systems, content-based multimedia retrieval databases make it difficult for a user to ask for information in a direct, precise query. Typically, content-based retrieval systems allow users to ask for information using examples of objects similar to the ones they wish to retrieve. Such an interface, however, requires mechanisms for the system to learn the query representation from the examples provided by the user. In our previous work, we proposed a query refinement mechanism in which a query representation is modified by the addition of new relevant examples based on user feedback. In this paper, we describe query processing mechanisms that can efficiently support query expansion via multidimensional index structure.

Porkaew, K., Mehrotra, S., & Winkler, R. (2000)
Database support for efficient visualization

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 121–125

> Effective visualization of information requires efficient techniques to support spatio-temporal queries over large terrain databases. This paper concentrates on continuous queries that correspond to a virtual mobile object visualizing the motions of other mobile objects in a dynamic environment. Continuous queries arise naturally during "fly-through" of a 3-D visualization. A naive approach to evaluating such queries is to repeatedly submit one query per visualized frame to the database. Since subsequent queries have a high degree of overlap with previous ones (because of the continuity of motion), a large amount of computation will be wasted. This paper proposes two alternate mechanisms that take

advantage of the continuity in the sequence of queries in order to optimize the evaluation cost of the queries.

Porkaew, K., Mehrotra, S., & Yu, H. (1999)
Continuous query in moving object databases to support efficient visualizations
Tech. Rep. No. TR-MARS-99-13, University of California, Irvine

Increasingly, application domains require database management systems to represent mobile objects and to support motion-specific queries. An important type of query in such domains is a continuous query, which consists of a sequence of instantaneous queries, one for each point of time $t' > t$, where t is the time the query is initially posed to the database. An example of a continuous query is monitoring objects within a specified distance of an object x, which itself may be mobile, starting at a given time t. A naive approach to evaluating continuous queries is to repeatedly submit instantaneous queries to the database, once for each point of time $t' > t$. Since subsequent queries have a high degree of overlap with the previous ones (because of the continuity of motion), much computation is wasted. This paper proposes two alternate mechanisms that attempt to reuse the answers returned by previous queries in evaluating subsequent queries, thereby optimizing the evaluation of continuous queries. Experiments conducted over a real-life dataset consisting of mobile objects (AHAS data containing Army battle exercises) are used to validate the efficiency of the developed approaches.

Pringle, H. L., Kramer, A. F., & Irwin, D. E. (2000)
Factors involved in perceptual change detection
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 33–36

The ability to detect changes in scenes can be used as a means to investigate how details in the world are perceived and remembered. Recent evidence indicates that changes in complex scenes are not readily detected, suggesting that individuals do not have a detailed, internalized representation of the world. The purpose of the research presented here is to examine the factors that may be important for predicting individual performance on a perceptual change detection task.

Pringle, H. L., Kramer, A. F., Irwin, D. E., Atchley, P. (1999)
Detecting changes in real-world scenes: The role of change characteristics and individual differences in attention
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 121–125

Recent research suggests that humans are surprisingly poor at detecting changes in scenes that occur during the course of eye movements. Indeed, this research has indicated that even large and apparently salient changes in scenes take a substantial amount of time to detect. In the present research, we examine the influence of several change characteristics (i.e.,

salience, meaning, and eccentricity) and individual differences in visual attention (i.e., the useful field of view) on perceptual change detection in the context of detailed driving scenes. These data are discussed in terms of how displays might be designed to help users to rapidly and accurately detect task-relevant changes.

Qian, R. J., & Huang, T. S. (1997)
Object detection using hierarchical MRF and MAP estimation
Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 186–192

This paper presents a new scale-, position-, and orientation-invariant approach to object detection. The proposed method first chooses attention regions in an image, based on the region detection result on the image. Within the attention regions, the method then detects targets via a novel object detection algorithm that combines template-matching methods with feature-based methods via hierarchical Markov random fields (MRFs) and maximum *a posteriori* probability (MAP) estimation. Hierarchical MRF and MAP estimation provides a flexible framework to incorporate various visual clues. The combination of template matching and feature detection helps to achieve robustness against complex backgrounds and partial occlusions in object detection. Experimental results are given in the paper.

Raghavan, V., & Molineros, J. (1999, June)
Interactive evaluation of assembly sequences using augmented reality
IEEE Transactions on Robotics and Automation, **15**, 435–449

This paper describes an interactive tool for evaluating assembly sequences via the novel human-computer interface of augmented reality. The goal is to enable the user to consider various sequencing alternatives of the manufacturing design process by manipulating both virtual and real prototype components. The augmented reality-based assembly evaluation tool would allow a manufacturing engineer to interact with the assembly planner while manipulating the real and virtual prototype components in an assembly environment. Information from the assembly planner can be displayed, superimposed directly upon the real. A sensing technique is proposed that uses computer vision along with a system of markers for automatically monitoring the assembly state as the user manipulates the assembly components. An implemented system called AREAS (augmented reality system for evaluating assembly sequences) is described. Also discussed is the advantage of using mixed prototyping and augmented reality as a means of capturing human intuition in assembly planning.

Rozenblit, J. W., Nugyen, H., & Barnes, M. J. (1999)
Effects of computer displayed color characteristics on individuals

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 163

The Advanced Battlefield Architecture for Tactical Information Selection (ABATIS) is described. ABATIS is a means of presenting battlefield information that facilitates understanding the process of the battle rather than simply the current location of various forces. The design of this system would reflect how the user assimilates battlefield-state information into a process-centered viewpoint. A key concept in the design of ABATIS is the process-centered display (PCD), a construct that can display complex, evolutionary processes, as well as simple, repetitive changes. For PCD to be effective, its architecture must support dynamic change, since battlefield processes (e.g., maneuver, attack) evolve and change as the battle unfolds, and must be flexible enough to permit the quick creation of new battle space objects from old ones. A secondary goal would be to use motion, color changes, morphing, or other types of animation to convey information. Some uses of animation are obvious, such as moving a symbol from one location to another. However, abstract quantities can also be tied to motion. A simple example would be representing the strength of a ground force by the speed of rotation of its symbol. When representation matches the intuitive notions of the user, the result is a metaphor that correlates familiar experiences with the actions of symbols.

Rudmann, D. S., Kramer, A. F., Bargar, R., Brady, R., & McCarley, J. (2000)
Cross-modal links in speech comprehension

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 179

A listener is better able to attend to one of two people speaking simultaneously if he can see as well as hear the speaker, especially if the sound is separated in space by being played over a loudspeaker at some distance from the person or from the televised image of the person speaking. We investigated this phenomenon, known as the ventriloquism effect, in realistic settings in which the number people speaking varies. Also, an eye tracker was used to determine the facial cues relied upon by a listener to comprehend the speech.

Rudmann, D.S., & McConkie, G. W. (April 30-May 2, 1998)
Acquiring spatial knowledge under varying field of view sizes

paper presented at the Midwestern Psychological Association Seventieth Annual Meeting, Chicago, IL

No abstract available

Rudmann, D. S., & McConkie, G. W. (1999)
Eye movements in human-computer interaction

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 91–95

> The potential benefits of incorporating eye movements into the interaction between humans and computers are numerous. For example, knowing the location of a user's gaze may help a computer to interpret a user's request, aid natural language processing, increase interaction by allowing the eyes to serve as a pointing device, and possibly enable a computer to ascertain some cognitive states of the user, such as confusion or fatigue. This paper details the problems encountered in previous attempts to use eye movements in human-computer interaction and evaluates current technology for its ability to overcome these limitations. An assessment of the accuracy and reliability of the ISCAN eye-tracking system (manufactured by Iscan, Inc.) and the pcBird head tracker (manufactured by Ascension Technology) is provided for two-dimensional displays. Recommendations are made for the design of eye-controlled display systems based on these technologies.

Rui, Y., Huang, T. S., & Chang, S.-F. (1998)
Digital image/video library and MPEG-7: Standardization and research issues

Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal
Processing, 6, 3785-3788

> Much research activity and interest has emerged in two closely related areas: the digital image/video library (DIVL) and MPEG-7. We review the critical research issues in DIVL from a signal processing viewpoint, the objectives and scope of MPEG-7, and the relationships between these two.

Rui, Y., Huang, T. S., & Mehrotra, S. (1998)
Browsing and retrieving video content in a unified framework

IEEE Second Workshop on Multimedia Signal Processing, 9–14

> In this paper, we first review the recent research progress in video analysis, representation, browsing, and retrieval. Motivated by the standard mechanisms for accessing book content (i.e., tables of contents and indexes), we then present novel techniques for accessing video content by constructing video equivalents. We further explore the relationship between video browsing and retrieval and propose a unified framework to seamlessly incorporate both entities. Preliminary research results justify our proposed framework for providing access to videos based on their content.

Rui, Y., Huang, T. S., & Mehrotra S. (1998)
Content-based image retrieval with relevance feedback in MARS

Proceedings of the IEEE International Conference on Image Processing, 2, 815–818

> Technological advances in the areas of image processing (IP) and information retrieval (IR) have evolved separately for a long time.

However, efficient content-based image retrieval systems require the integration of the two. We attempted to link the image retrieval model to the text retrieval model, so that the well-established text retrieval techniques can be used to retrieve relevant imagery. Specifically, we propose an approach of mapping the image feature vector (IP domain) to weighted term vector (IR domain). The relevance feedback technique from the IR domain is used to demonstrate the effectiveness of this mapping. Experimental results show that the image retrieval precision increases considerably with the help of relevance feedback.

Rui, Y., Huang, T. S., & Mehrotra, S. (1998)
Exploring video structure beyond the shots
Proceedings of IEEE International Conference on Multimedia Computing and Systems, 237–240

While existing shot-based video analysis approaches provide users with better access to the video than do raw data streams, they are still not sufficient for meaningful video browsing and retrieval, since (1) the shots in a long video are still too numerous to be presented to the user, and (2) shots do not capture the underlying semantic structure of the video, the basis upon which the user may wish to browse/retrieve the video. To explore video structure at the semantic level, this paper presents an effective approach to extracting the underlying video scene structure and grouping shots into semantically related scenes. The output of the proposed algorithm provides a structured video that greatly facilitates the user's access. Experiments based on real-world movie videos validate the effectiveness of the proposed approach.

Rui, Y., Huang, T. S., & Mehrotra S. (1998, January)
Human perception subjectivity and relevance feedback in multimedia information retrieval
Paper presented at the meeting of the IS&T and SPIE Storage and Retrieval for Image and Video Databases IV, San Jose, CA

Content-based multimedia information retrieval (MIR) has become one of the most active research areas in the past few years. While the existing approaches establish the basis of MIR, techniques for incorporating the subjectivity of human perception into the retrieval process have not been fully investigated. To address this problem, this paper introduces an integrated relevance feedback architecture for MIR, which dynamically captures the user's perception subjectivity and simultaneously models it at various levels by using dynamically updated weights based on the user's relevance feedback. The experimental results show that the proposed approach greatly reduces the user's effort of composing a query and captures the user's information needs more precisely.

Rui, Y., Huang, T. S., Mehrotra, S., & Ortega, M. (1997)
Automatic matching tool selection via relevance feedback in MARS

The 2nd International Conference on Visual Information Systems, 109–116

Because of the diversity in subjective human judgment, a visual information retrieval system that supports a single prefixed similarity measure will result in poor retrieval performance. To address this problem, this paper proposes the concept of a similarity matching toolkit, which consists of different similarity measures that simulate human perception of a given feature from different aspects. The toolkit supports a feedback-driven tool-selection mechanism that adapts to the similarity measure that best fits the user's perception. To illustrate the advantage of the proposed toolkit approach, we apply it to shape-based image retrieval. We describe a shape-matching toolkit, which consists of four transformation-invariant and computationally efficient matching tools, and describe how relevance feedback can be used for automatic tool selection. Experimental results validate the flexibility of the matching toolkit and show the effectiveness of the relevance feedback for shape-matching tool selection.

Rui, Y., Huang, T. S., Mehrotra, S., & Ortega-Binderberger, M. (1997)
A relevance feedback architecture for content-based multimedia information retrieval systems

Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Librarie, 82–89

Content-based multimedia information retrieval (MIR) has become one of the most active research areas in the past few yeas. Many retrieval approaches based on extracting and representing visual properties of multimedia data have been developed. While these approaches establish the viability of MIR based on visual features, techniques for incorporating human expertise to improve retrieval performance have not been studied. To address this limitation, this paper introduces a human-computer interaction-based approach to MIR, in which the user guides the system during retrieval using relevance feedback. Our experiments show that the retrieval performance is significantly improved by incorporating humans in the retrieval process.

Rui, Y., Huang, T. S., Ortega, M., & Mehrotra, S. (1998)
Relevance feedback: A power tool for interactive content-based image retrieval

IEEE Transactions on Circuits and Systems for Video Technology, 8, 644–655

Content-based image retrieval (CBIR) has become a highly active research area in the past few years. Many visual feature representations have been explored and many systems built. While these research efforts establish the basis of CBIR, the usefulness of the proposed approaches is limited. Specifically, these efforts have generally ignored two distinct characteristics of CBIR systems: (1) the gap between high-level concepts

and low-level features, and (2) the subjectivity of human perception of visual content. This paper proposes an interactive retrieval approach, based on relevance feedback, which effectively takes into account these two characteristics in CBIR. During the retrieval process, the user's high-level query and perception subjectivity are captured by dynamically updated weights based on the user's feedback. Experimental results for more than 70,000 images show that the proposed approach greatly reduces the user's effort in composing a query and captures the user's information need more precisely.

Schlabach, J. L., Goldberg, D. L., Hayes, C. C. (1999)
FOX-GA: A genetic algorithm for generating and analyzing battlefield courses of action
Evolutionary Computation, 7, 45–68

This paper describes Fox-GA, a genetic algorithm (GA) that generates and evaluates plans in the complex domain of military maneuver planning. Fox-GA's contributions are to demonstrate an effective application of GA technology to a complex, real-world planning problem and to provide an understanding of the properties needed in a GA solution to meet the challenges of decision support in complex domains. Previous obstacles to applying GA technology to maneuver planning include the lack of efficient algorithms for determining the fitness of plans. Detailed simulations would ideally be used to evaluate these plans, but most such simulations typically require several hours to assess a single plan. Since a GA needs to quickly generate and evaluate thousands of plans, these methods are too slow. To solve this problem, we developed an efficient evaluator (wargamer) that uses coarse-grained representations of this problem domain to allow appropriate yet intelligent trade-offs between computational efficiency and accuracy. An additional challenge was that users needed a set of significantly different plan options from which to choose. Typical GAs tend to develop a group of "best" solutions that may be very similar (or identical) to each other. This may not provide users with sufficient choice. We addressed this problem by adding a niching strategy to the selection mechanism to ensure diversity in the solution set, providing users with a more satisfactory range of choices. Fox-GA's impact will be in providing decision support to constrained and cognitively overloaded battle staff to help them rapidly explore options, create plans, and better cope with the information demands of modern warfare.

Schlabach, J. L., & Hayes, C. C. (1998)
Fox-GA: A genetic algorithm for generating and analyzing courses of action

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 39–43

> This paper describes Fox-GA, a genetic algorithm (GA) that generates and evaluates battlefield courses of action (COAs). A previous obstacle to applying GA technology to COA evaluation was the lack of efficient algorithms for determining the fitness of COAs. Detailed simulations are typically used to evaluate COAs, but they typically require several hours for a each COA. Since GAs need to quickly generate and evaluate thousands of COAs, detailed simulations are too slow. To solve the problem, we developed an efficient evaluator (wargamer) that uses coarse-grained representations to allow efficient assessments of COAs.

Servetto, S., Ramchandran, K., & Huang, T. S. (1997)
A successively refinable wavelet-based representation for content-based image retrieval

1997 IEEE First Workshop on Multimedia Signal Processing, 325–330

> Content-based retrieval of image and video data from databases is a very challenging problem that must be solved to support efficient access to vast amounts of visual information. Typical queries to be performed in this context check attributes of objects present in image data, such as shape, color, relative locations, and so forth. Therefore, the way in which image data are represented plays a fundamental role in the efficient implementation of those queries. One possibility is to take the inefficient approach of storing images via standard compression techniques, storing image features (such as object shape descriptors, color histograms, etc.) as explicit side information, and whenever an image is involved in the evaluation of a query, decoding it to full resolution. However, more efficient techniques (in terms of storage and computational requirements) are possible. We propose a new image coding technique (which combines a wavelet image representation, embedded coding of the wavelet coefficients, and segmentation of semantically meaningful objects in the wavelet domain) to generate a bit stream in which each object is encoded independently of every other object in the image, without the need for explicitly storing shape boundary information. Since the representation of each object is fully embedded, applications may specify, independently for each object, the desired target bit rate and may retrieve bits from the compressed bit stream.

Servetto, S. D., Ramchandran, K., & Huang, T. S. (1998)
Image and video coding with object indexing support

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 96–100

> We propose new coding techniques that combine a wavelet
> representation, embedded coding of the wavelet coefficients, and
> segmentation of semantically meaningful objects in the wavelet domain
> to generate a bit stream in which semantically meaningful, arbitrarily
> shaped image objects are encoded independently of each other, without
> the need for explicitly storing shape boundary information. Since the
> representation of each object is fully embedded, applications may specify,
> independently for each object, the desired target bit rate and retrieve bits
> from the compressed bit stream. Simulation results show that these new
> proposed indexing methods achieve coding performance which is
> perceptually identical to that achieved via state-of-the-art image/video
> coding techniques that do not support indexing, thus proving the
> feasibility of generating bit streams that can support functionality
> required by emerging multimedia applications without sacrificing
> compression performance.

Servetto, S. D., Ramchandran, K., & Orchard M. T. (1997, September)
Wavelet-based image coding via morphological prediction of significance

Paper presented at the International Conference on Image Processing, Santa
Barbara, CA

> In previous work, we introduced a new image representation for the field
> of wavelet coefficients (dubbed morphological representation of wavelet
> data [MRWD]), based on morphological operators. The present work
> extends the MRWD framework by addressing the effective design of
> image-coding algorithms. First, we design an encoder with the goal of
> being optimal in the operational rate-distortion sense. Second, based on
> the same (morphological) techniques, we design a successively refinable
> version of the single rate coder. Finally, we report simulation results
> using these techniques.

Servetto, S. D., Ramchandran, K., & Orchard, M. T. (1999)
Image coding based on a morphological representation of wavelet data
IEEE Transactions on Image Processing, 8, 1161–1174

> An experimental study of the statistical properties of wavelet coefficients
> of image data is presented, as well as the design of two different
> morphology-based image-coding algorithms that use these statistics. A
> salient feature of the proposed methods is that, by a simple change of
> quantizers, the same basic algorithm yields high-performance embedded
> or fixed rate coders. Another important feature is that the shape
> information of morphological sets used in this coder is encoded implicitly
> by the values of wavelet coefficients, thus avoiding the use of explicit and
> rate-expensive shape descriptors. These proposed algorithms, while

achieving nearly the same objective performance as state-of-the-art zero-tree-based methods, can produce reconstructions of a somewhat superior perceptual quality because they exhibit a property of compression and noise reduction.

Servetto, S., Rosenblatt, J., & Ramchandran, K. (1997)
A binary Markov model for the quantized wavelet coefficients of images and its rate/distortion optimization

International Conference on Image Processing, **3**, 82–85

Zero-tree-based algorithms represent the state of the art in wavelet-based image coding. These algorithms can be generally described as first sending some map of locations of zero coefficients (the set of zero-tree symbols) and then sending the value of non-zero coefficients. However, the decision of what map to send is typically made with some simplifying assumption about the structure of the map, motivated by some empirically observed property of the data (e.g., that zero coefficients are likely to appear in tree-structured sets). In the present paper, the map of locations of zero coefficients is optimally estimated as a hidden binary Markov random field. Algorithms are presented for estimating the hidden field, given the observed wavelet coefficients, for encoding the field, and for encoding the data given the field estimate. Simulation results show the coding algorithm to possesses a rate-distortion performance that is equal or superior to any published zero-tree-based image coder. This fact provides conclusive empirical evidence that the proposed model is appropriate for the data.

Servetto, S. D., Rui, Y., Ramchandran, K., & Huang, T. S. (1999)
A region-based representation of images in MARS

Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology, **20**, 137–150

We study the problem of representing images within a multimedia database management system to support fast retrieval operations without compromising storage efficiency. To achieve this goal, we propose new image-coding techniques that combine a wavelet representation, embedded coding of the wavelet coefficients, and segmentation of image-domain regions in the wavelet domain. A bit stream is generated in which each image region is encoded independently of other regions, without the need to store information describing the regions. Simulation results show that our proposed algorithms achieve coding performance that compares favorably, both perceptually and objectively, to that achieved by state-of-the-art image/video coding techniques, while additionally providing region-based support.

Sethi, Y. (1998)
Multimodal analysis of gesture and speech in video sequences
Unpublished master's thesis, The Pennsylvania State University, University Park, PA

> A gesture recognition system, based on hidden Markov modeling, was developed to make possible machine recognition of gestures and thus enable more natural human-computer interaction. A hidden Markov model was "trained" to recognize a few natural gestures produced by weather reporters during weather forecasts and then validated on television weathercasts. Gesture recognition was highly accurate (100%), with discrete gestures isolated from a continuous stream of gestures, but less accurate (about 56%) when the targeted gestures were part of a stream of movements. Accuracy for streamed gestures increased (by about 12%) when a speech recognition system was used in combination with the gesture-recognition system to detect words that co-occurred with the targeted gestures.

Sharma, R., Huang, T. S., & Pavlovic V. I. (1996)
A multimodal framework for interacting with virtual environments
In C. A. Ntuen & E. H. Park (Eds.), *Human interaction with complex systems: Conceptual principles and design practice* (pp 53–71). Boston: Kluwer Academic Publishers

> Although there has been a tremendous progress in recent years in three-dimensional, immersive display, and virtual reality (VR) technologies, the corresponding interface technologies have lagged behind. To fully exploit the potential that VR offers as a means of visualizing and interacting with complex information, it is important to develop "natural" means for interacting with the virtual display. Such natural interaction can be achieved by an integrated approach in which multiple, possibly redundant modes of input such as speech, hand gesture, gaze, and graphical feedback are used simultaneously. This paper presents a conceptual framework for multimodal human-computer interaction for manipulating a virtual object. Specific techniques are presented for using a combination of speech and gesture for manipulating virtual objects. Free-hand gestures are analyzed and recognized by computer vision. The gesture analysis is done cooperatively with the speech recognition system and the graphic system. This is demonstrated with the help of an experimental VR setup used by molecular biologists for simulating and visualizing complex molecular structures.

Sharma, R., & Hutchinson, S. (1997)
Motion perceptibility and its application too active vision-based servo control
IEEE Transactions on Robotics and Automation, **13**, 607–617

> We address the ability of a computer vision system to perceive the motion of an object (possibly a robot manipulator) in its field of view. We derive a quantitative measure of motion perceptibility, which relates the

magnitude of the rate of change in an object's position to the magnitude of the rate of change in the image of that object. We then show how motion perceptibility can be combined with the traditional notion of manipulability into a composite perceptibility/manipulability measure. We demonstrate how this composite measure can be applied to a number of different problems involving relative hand-eye positioning and control.

Sharma, R., Pavlovic, V., & Huang, T. S. (1998)
Toward multimodal human computer interaction
Proceedings of the IEEE, 86, 853–869

> Recent advances in various signal processing technologies, coupled with an explosion in available computing power, have given rise to a number of novel human-computer interaction (HCI) modalities: speech, vision-based gesture recognition, eye tracking, electroencephalograph, and so forth. Successful incorporation of these modalities into an interface could potentially ease the HCI bottleneck that has become noticeable with the advances in computing and communication. It has also become increasingly evident that the difficulties encountered in the analysis and interpretation of individual sensing modalities may be overcome by their integration into a multimodal human-computer interface. We examine several promising approaches to achieving multimodal HCI. We consider some of the emerging novel input modalities for HCI and the fundamental issues in integrating them at various levels, from early signal level to intermediate feature level to late decision level. We discuss the different computational approaches that may be applied at the different levels of modality integration. We also briefly review several demonstrated multimodal HCI systems and applications. Despite all the recent developments, it is clear that further research is needed for interpreting and fitting multiple sensing modalities in the context of HCI. This research can benefit from many disparate fields of study that increase our understanding of the different human communication modalities and their potential role in HCI.

Sharma, R., Poddar, I., & Kettebekov, S. (2000)
Recognition of natural gestures for multimodal interactive map (iMAP)
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 77–81

> Previous attempts at incorporating gesture recognition in multimodal human-computer interaction (HCI) have resulted in systems that use predefined gestures. However, recognition of predefined gestures is not conducive to developing "natural" HCI. In the present paper, we propose a novel approach for continuous recognition of deictic natural gestures. Deictic natural gestures occur in combination with speech and can be used extensively in the context of HCI. These gestures are not predefined and they do not string together in syntactic bindings like sign language. The proposed approach uses a real-time vision-based predictive feature-tracking algorithm. The classification and continuous recognition of

gestures is based on hidden Markov models. Gesture recognition is integrated into a multimodal framework that allows a user to interact naturally with a graphical display using speech and gesture. We describe the evolution of the experimental test bed, iMAP, that allows free-hand gestures and spoken words for interacting with a campus map. Extensive tests using this framework obtained natural gesture recognition rates as high as 80%. However, multimodal integration and constraints from system interpretation are necessary for further improvement of accuracy. We use the context of an interactive campus map to discuss the critical components of the multimodal interpretation and integration problems.

Sharma, R., Poddar, I., Ozyildiz, E., Kettebekov, S., Kim, H., & Huang, T. S. (1999)
Toward interpretation of natural speech/gesture: Spatial planning on a virtual map

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 35–39

Hand gestures and speech are the most important modalities of human-to-human interaction. Accordingly, there is considerable interest in incorporating these modalities into "natural" human-computer interaction (HCI), particularly within virtual environments. An important feature of such a natural interface would be an absence of predefined speech and gesture commands. The resulting bimodal speech-gesture HCI "language" would thus have to be interpreted by the computer. While some progress has been made in the natural language processing of speech, the inclusion of gestures is even more challenging. This challenge ranges from the low-level signal processing of bimodal (audio/video) input to the high-level semantic interpretation of natural speech/gesture. In this paper, we consider the design of a speech-gesture interface in the context of a set of spatial tasks defined on a virtual map of an urban area. The task constraints then make it feasible to study the critical components of the bimodal interpretation problem and define an agent-based architecture for implementing the interface. An experimental test bed is also described, in which free-hand gestures and spoken words are used for spatial planning tasks defined on a virtual two-dimensional map. Such tasks would also be involved in crisis management, mission planning, and briefing.

Sharma, R. & Sutanto, H. (1997)
Integrating configuration space and sensor space for vision-based robot motion planning

In J.-P. Laumond & M. Overmars (Eds.), *Algorithmic foundations of robotics* (pp 63–78). Wellesley, MA: A. K. Peters

Visual feedback can play a crucial role in a dynamic robotic task such as the interception of a moving target. To use the feedback effectively, there is a need to develop robot motion-planning techniques that also take into account properties of the sensed data. We propose a motion-planning

framework that achieves this with the help of a space called the perceptual control manifold (PCM), defined on the product of the robot configuration space and an image-based feature space. We show how the task of intercepting a moving target can be mapped to the PCM, using image feature trajectories of the robot end effector and the moving target. This leads to the generation of motion plans that satisfy various constraints and optionality criteria derived from the robot kinematics, a control system, and the sensing mechanism. Specific interception tasks are analyzed to illustrate this vision-based planning technique.

Shattuck, L., Graham, J., Merlo, J., & Hah, S. (2000)
Cognitive integration: An investigation of how expert and novice commanders process battlefield data
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 47–51

Technology provides military decision makers with more data than they can possibly use. Commanders and staffs must sort through and combine relevant data to develop understanding. This process, which we call cognitive integration, was investigated in a tactical simulation using 21 experienced active duty Army officers (former battalion commanders) and 21 novice officers (no battalion command experience) as participants. Quantitative and qualitative data yielded significant differences between the experienced and novice groups. In addition, data analysis led to the development of several important design principles that will be used to build a decision aid prototype to assist commanders in integrating data.

Sistla, A. P., Wolfson, O., Chamberlain, S., & Dao, S. (1998)
Querying the uncertain position of moving objects
in O. Etizon, S. Jajodia, S. Sripada (Eds.), *Temporal Databases: Research and Practice* (pp 310–337). Berlin, Germany: Springer-Verlag

The authors propose a data model for representing moving objects with uncertain positions in database systems: the Moving Objects Spatio-Temporal (MOST) data model. They also propose Future Temporal Logic (FTL) as the query language for the MOST model and devise an algorithm for processing FTL queries in MOST.

Sistla, A. P., Wolfson O., & Huang, Y. (1998)
Minimization of communication cost through caching in mobile environments
IEEE Transactions on Parallel and Distributed Systems, 9, 378–390

Users of mobile computers will soon have on-line access to a large number of databases via wireless networks. Because of limited bandwidth, wireless communication is more expensive than wire communication. In this paper, we present and analyze various static and dynamic data allocation methods. The objective is to minimize the communication cost between a mobile computer and the stationary computer that stores the on-line database. Analysis is performed on two

cost models. One is connection (or time) based (as in cellular telephones), where the user is charged per minute of connection. The other is message based (as in packet radio networks), where the user is charged per message. Our analysis addresses both the average case and the worst case for determining the best allocation method.

Sistla, A. P., Wolfson, O., Yesha, Y., & Sloan, R. H. (1998)
Towards a theory of cost management for digital libraries and electronic commerce
ACM Transactions on Database Systems, 23, 411–452

> One feature that distinguishes digital libraries from traditional databases is new cost models for client access to intellectual property. Clients will pay to access data items in digital libraries, and we believe that optimizing these costs will be as important as optimizing performance for traditional databases. We discuss cost models and protocols for accessing digital libraries, with the objective of determining the minimum cost protocol for each model. We expect that in the future, information appliances will come equipped with a cost optimizer, in the same way that computers today come with a built-in operating system. We make the initial steps toward a theory and practice of intellectual property cost management.

Sniezek, J. A., & Chernyshenko, O. S. (1999)
Psychological evaluation of Co-RAVEN technology for battlefield decision making: Probabilistic reasoning by Army intelligence experts
Tech. Rep. No. 99-1, Urbana-Champaign, IL: University of Illinois, Department of Psychology

> Co-RAVEN is a Bayesian-based decision aid that generates probabilities of the occurrence of high-level events from detailed data. The Bayesian decision net is created by the encoding of probability statements from actual military intelligence experts on real-world intelligence problems. However, a common finding of decision research is that decision makers are often overconfident in their judgments. This paper found that intelligence officers exhibit overconfidence in their decisions and that they do not agree in their probability estimates. The implications of these findings for creating Bayesian decision nets was discussed.

Sniezek, J. A., & Schrah, G. E. (2000)
Effects of communication medium, judge-advisor roles, and information load on decision processes
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 43–45

> Changes in the information environment have changed the way in which people make mission-critical decisions. Teams consisting of individuals with different expertise and roles have replaced the individual decision maker, and the sources of information available to them have increased dramatically. The Judge-Advisor System (JAS) acts as a model of such

teams and is used in this experiment to investigate the influence of communication medium, group role, and information load on decision processes. Results indicate that information load has predictable effects on decision processes and ultimately on performance. These results have substantial implications for decision making in information-rich environments.

Srinivasa, N., & Sharma R. (1997)
Execution of saccades for active vision using a neuro-controller

IEEE Control Systems Magazine, **17**, 18–29

An important mechanism in active vision is fixating on different targets of interest in a scene. We propose a two-stage execution of saccades, in which the first stage is an "open loop" mode based on a learned spatial representation, and the second stage is a closed loop "visual serving" mode. Explicit calibration of the kinematic and imaging parameters of the system is replaced with a self-organized learning scheme, thereby providing a flexible and efficient saccade control strategy. Experiments on the University of Illinois Active Vision System (UIAVS) are used to establish the feasibility of this approach.

Stroming, J. W., Kang, Y., Huang, T. S., & Kang, S. M. (1997)
New architectures for modified MMR shape coding

Proceedings of the 1997 IEEE International Symposium on Circuits and Systems: Circuits and Systems in the Information Age, **2**, 1205–1208

New architectures for modified MMR shape encoding and decoding are presented. MPEG-4 is first briefly described, as is the modified MMR algorithm proposed for use in MPEG-4 shape coding. Architectures for encoding and decoding, which reduce memory access, use custom hardware to accelerate critical components, require little external control, and can be easily pipelined or parallelized, are described.

Sundareswaran, V., & Behringer, R. (1998)
Virtual interaction with real object displays

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 59–63

We present an augmented reality system based on a dynamic tracking procedure. We affix markers to areas to be registered and we track the markers on the image. We use two-dimensional screen coordinates of the markers and known three-dimensional configuration of the markers to adjust the position and orientation of the camera in the virtual scene, resulting in registration of the graphical model with the real object. Based on this registration procedure, we have created a display interface that allows users to interact virtually with real object displays. Examples of such interactions include "x-ray" vision capability, passive wireless interfacing, and introduction of virtual objects in real scenes. In this paper, we present an overview of our system and focus on its interface capabilities.

Sundareswaran, V., & Behringer, R. (1999)
Visual servoing-based augmented reality

in Behringer, R., Klinker, G., & Mizell, D. W. (Eds.). *Augmented reality: Placing artificial objects in real scenes* (pp 193–200). Natick, MA: A. K. Peters

> One of the central problems of augmented reality (AR) is the accurate placement of virtual objects in the real world. This paper focused on video-based AR, in which a camera is used to generate an image of the real world; the image is then processed to determine where the computer-generated elements should be displayed. To accomplish this, the relative position and orientation of the camera must be known.

Sundareswaran, V., & Chen, S. (1999)
Hand-held displays for control and communication with large format displays

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 165

> In a demonstration, a hand-held personal computer (HPC) was used to control the view on a large display. The large display on a desktop computer (Windows™ 95, DirectX) showed only a portion of a pre-rendered isometric view of a battlefield. Animated units were controlled through a stylus, a graphics tablet, and the HPC. Troop movement and identified red unit positions were displayed on both the large and the HPC displays. Control was achieved through stylus interaction and speech commands in a multimodal fashion. Control of wireless integrated network sensors is demonstrated, as well as a display of the situation reported by the sensors.

Sundareswaran, V. S., Chen, S. L., McGee, J. H., & Vassiliou, M. (2001)
An integrated displays testbed for multi-modal interaction

Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 111–115

> An important component in prototyping new multi-modal human-computer interaction (HCI) technologies is a means for testing the technologies in an integrated framework. The framework should accommodate of a variety of platforms, possess a simplified architecture, and be easily extendible. To address these requirements, we have developed an HCI test bed under the FedLab program. This test bed contains features that meet the above requirements. In this paper, we provide a brief description of the test bed, representative test bed components, and user interface elements.

Sutanto, H., Sharma, R., & Varma, V. (1997)
Image-based autodocking without calibration

Proceedings of the 1997 IEEE International Conference on Robotics and Automation, 974–979

> The calibration requirements for visual servoing can make it difficult to apply in many real-world situations. One approach to image-based visual servoing without calibration is to dynamically estimate the matrix (called the image Jacobian) that relates changes in the robot effectors to changes in the image and use the image Jacobian as the basis for control. However, with the normal motion of a robot toward the goal, the estimation of the image Jacobian deteriorates over time. We propose the use of additional "exploratory motion" to considerably improve the estimation of the image Jacobian. We study the role of such exploratory motion in a visual servoing task. Simulations and experiments with a robot possessing six degrees of freedom are used to verify the practical feasibility of the approach.

Tang, H., & Beebe, D. J. (1999)
Tactile sensitivity of the tongue on photo-lithographically fabricated patterns

Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 167

> Previous psychology and neuroscience studies suggest that the oral cavity is a sensory-rich location. Recent advances in miniaturization technologies make it possible to build tactile devices that can operate within the oral cavity. In order to design optimal tactile interfaces for the mouth, we must understand the perceptual characteristics of the mouth. This report describes preliminary work aimed at measuring several perceptual parameters of the tongue's tip and anterior dorsal surface.

Tang, H., & Beebe, D.J. (1999)
An ultra-flexible electrotactile display for the roof of the mouth

Proceedings of the First Joint BMES/EMBS Conference, 1, 626

> No abstract available

Tang, H., Beebe, D. J., & Kramer, A. F. (1999)
Comparison of tactile and visual feedback for a multi-state input mechanism

Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Magnificent Milestones and Emerging Opportunities in Medical Engineering, 4, 1697–1700

> A chording system that incorporates pressure-sensitive input elements and vibratory feedback elements is presented. Both input and feedback elements are capable of multiple states. A three-state, three-element system, in which input states correspond to levels of finger pressure on sensor, was used for the experiment. Feedback was provided tactilely by

stimulators vibrating in bursts on the palm and/or visually as colors on a screen, creating three treatment conditions: visual, tactile and combined visual-tactile feedback. The influence of feedback on the speed and accuracy with which the subjects could input information was examined. The preliminary results indicate within-modality tactile feedback provides improved performance over cross-modality feedback or simultaneous visual and tactile feedback.

Tang, L., & Huang, T. S. (1996)
Automatic construction of 3D human face models based on 2D images
Proceedings of International Conference on Image Processing, 3, 467–470

> We propose an approach to the automatic construction of three-dimensional human face models using a generic face model and several two-dimensional images of an actual human face. A template-matching algorithm automatically extracts all necessary facial features from the front and side profile of the images of a person's face and then fits the generic face model to these feature points by geometric transforms. Finally, texture mapping is performed to achieve realistic results.

Tao, H., & Huang, T. S. (1997)
Modeling spatial-temporal patterns in facial actuation
Proceedings of the IEEE CVPR'97 Non-rigid and Articulated Motion Workshop, 54–60

> In this paper, a new method of modeling human facial articulation is proposed. The approach contains three major parts: reducing the spatial dimension through principal component analysis; approximating the temporal function using a simple basis function, which is similar to facial articulation process; and improving recognition and compression capability by means of a learning algorithm. This scheme is also used for encoding facial articulation parameter sequences. Though developed based on the facial animation parameter set (MPEG-4 facial animation parameter set), the algorithm can be easily applied to other parameter representations.

Tao, H., & Huang, T. S. (1997)
Multi-scale image warping using weighted Voronoi diagram
Proceedings of the International Conference on Image Processing, 1, 241–244

> We propose a new multi-scale image warping method based on the weighted Voronoi diagram. Weights are assigned to the control points according to their influence scales. At each scale level, a triangulation is constructed, based on the weighted Voronoi diagram. Then the interpolation of displacements is performed on this triangulation. In this process, only the control points with large weights are mapped to their final values. Once the control points with large weights have been mapped correctly, their weights are modified to be the maximum weight of those unmapped control points, and their displacement values are set to 0. This process is performed iteratively until all control points are correctly mapped. The advantage of this approach is that the underlying

triangulation changes between scales to fit the warping scale. Both global warping and local warping can be modeled appropriately with this approach.

Tao, H. & Huang, T. S. (1998)
Bézier volume deformation model for facial animation and video tracking
In N. Magnenat-Thalmann and D. Thalmann (Eds.), *Modeling and Motion Capture Techniques for Virtual Environments* (pp 242–253). Berlin, Germany: Springer-Verlag

> Capturing real motions from video sequences is a powerful approach for automatically building a facial deformation model. In this paper, a three-dimensional Bézier volume deformation model is proposed for both synthesis and analysis of facial movements. Since this model is independent of the mesh structure (provided that the feature points are given), it can animate geometric facial models of different shapes and structures. Of equal importance, the linear property of this model implies a simple and robust analysis algorithm, from which a customized facial deformation model is derived. Experimental results of animation and video analysis are demonstrated.

Tao, H., & Huang, T. S. (1999)
Facial motion synthesis and analysis using a free-form deformation model
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 169

> Capturing real facial motions from video sequences is a powerful approach for automatic generation of a facial deformation model. In this paper, a three-dimensional piecewise Bézier volume deformation model is proposed for both facial animation and facial motion analysis. Because this model is independent of the mesh structure, the resulting deformation model can be used for animating different geometric face models. The more important linear property of this model also implies an efficient and robust analysis algorithm, from which a customized facial deformation model can be derived. Experimental results of facial animation and video analysis are demonstrated.

Tayeb, J., Ulusoy, O., & Wolfson, O. (1998)
A quadtree-based dynamic attribute indexing method
Computer Journal, 41, 185–200

> Dynamic attributes change continuously over time, making it impractical for explicit updates to be issued for every change. In this paper, we adapt a variant of the quadtree structure to solve the problem of indexing dynamic attributes. The approach is based on the key idea of using a linear function of time for each dynamic attribute so that we can predict its value in the future. We contribute an algorithm for regenerating the quadtree-based index periodically, which minimizes CPU and disk access costs. We also provide an experimental study of performance, focusing on query processing and index update overheads.

Theeuwes, J., Atchley, P., & Kramer, A. F. (2000)
I: Control of visual attention

In S. Monsell & J. Driver (Eds.), *Attention and Performance XVIII* (pp 71–208). Cambridge, MA: The MIT Press

> Previous research has shown that a salient feature singleton captured attention in a "bottom-up" fashion. A salient color singleton interfered with subjects' search for a less salient shape singleton, which suggests that early processing is driven by bottom-up saliency factors. In the present experiments, we examined how bottom-up and top-down processing develops over time. Subjects searching for a shape singleton target had to ignore a color singleton distractor, which was presented at different stimulus onset asynchronies before the search display. The results indicate that when the target and distractor were presented simultaneously, the salient singleton distractor captured attention. However, when the distractor singleton was presented about 150 milliseconds before the target singleton, the distractor did not disrupt performance. The findings suggest a stimulus-driven model of selection in which early processing is solely driven by bottom-up activation, but later, visual processing can resist the distractor in that it can be overridden by top-down attentional control.

Theeuwes, J., Kramer, A. F., & Atchley, P. (1998)
Attentional control within 3-D space

Journal of Experimental Psychology: Human Perception and Performance, **24,** 1476–1485

> Four experiments investigated whether directing attention to a particular plane in depth enables observers to filter out information from another depth plane. Observers searched for a red line segment among green line segments in stereoscopic displays. Results showed that directing attention to a particular depth plane cannot prevent attentional capture from another depth plane when the colors of the target and distractor are identical. However, attentional capture by a singleton from another depth plane is prevented when the colors of the target and distractor are different. These results indicate that only when both color and depth information are selective in guiding attention to the target singleton can attentional capture by irrelevant singletons be prevented. The results also suggest that retinal disparity does not have the same special status as location information in two dimensions and should be considered as just another feature along which selection may occur.

Theeuwes, J. Kramer A. F. & Atchley, P. (1998)
Visual marking of old objects

Psychonomic Bulletin and Review, 5, 130–134

> D.G. Watson and G.W. Humphreys presented evidence that selection of new elements can be prioritized by on-line, top-down attentional inhibition of old stimuli already in the visual field (visual marking). The

experiments on which this evidence was based always presented old elements in green and new elements in blue; selection could therefore have been based on color. The present experiment, which does not contain this confound, showed that visual marking is a strong and robust process that enables subjects to visually mark at least 15 old elements, even when these elements are the same color as the new ones. The results indicate that preview of the elements is critical—not the fact that those elements contained a common feature.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998)
Our eyes do not always go where we want them to go: Capture of the eyes by new objects
Psychological Science, 9, 379–385

Observers make rapid eye movements to examine the world around them. Before an eye movement is made, the observer's attention covertly shifts to the location of the object of interest. The eyes will typically land at the position at which attention is directed. Here, the authors report that a goal-directed eye movement toward a uniquely colored object is disrupted by the appearance of a new but task-irrelevant object unless subjects (n =15) have enough time to focus their attention on the location of the target before the appearance of the new object. In many instances, the eyes started moving toward the new object before gaze started to shift to the color-singleton target. The eyes often landed for a very short period of time (25 to 150 ms) near the new object. The results suggest parallel programming of two saccades: one voluntary, goal-directed eye movement toward the color-singleton target and one stimulus-driven eye movement reflexively elicited by the appearance of the new object. Neuro-anatomical structures responsible for parallel programming of saccades are discussed.

Theeuwes, J., Kramer, A. F., Hahn, S., & Waite, T. (1997, May)
Effect of the appearance of a new object on oculomotor control
Poster session presented at the meeting of Investigative Ophthalmology and Visual Science

Our purpose was to determine the extent to which endogenously controlled eye movements are affected by the sudden appearance of an irrelevant object (an abrupt onset). Previous research has shown that visual attention is captured by the appearance of a new object (e.g., Yantis, 1993). Observers had to make an eye movement to a predefined target present in the visual field. At different stimulus onset asynchronies (SOAs) after the presentation of the target (0, 50, 100, 150 milliseconds), an abrupt onset was presented at different locations in the visual field. Both manual and eye latencies were measured as well as the scan patch of the eye. The results indicate that at the early SOAs, latencies to respond to the target were increased when the onset was presented near the target location. At later SOAs and at locations away from the target, the onset

had no effect. The appearance of a new object in the visual field does provoke a preset eye movement toward the object when that new object is presented near the target. The results suggest that similar to visual attention, eye movement behavior is the result of an interaction between goal-driven and stimulus-driven factors.

Thomas, L. C., & Wickens, C. D. (1999)
Immersion and battlefield visualization: Frame of reference effects on navigation tasks and cognitive tunneling
Proceedings of the Human Factors & Ergonomics Society 43rd Annual Meeting, 153–157

> Army officers viewed a battle scenario in one of two computer-based display conditions. The tethered condition was a three-dimensional (3-D) exocentric display, and the immersed condition was a 3-D egocentric display, which allowed 360° panning, coupled with a small two-dimensional contour map of the entire battle area embedded in the top center of the screen. The participants' tasks included providing verbal reports of new enemy units or changes in existing units, responding to a series of diagnostic questions, and giving confidence ratings for their selected answers. Results showed that participants in the immersed condition were less accurate on questions that required panning, as well as on questions requiring a count of enemy units, than in the tethered condition. However, confidence ratings for both tasks did not differ between display conditions. These findings indicate that participants in the immersed condition "cognitively tunneled" into the initial forward field of view, relying too heavily on information in this view and not adequately panning the environment.

Thomas, L. C., & Wickens, C. D. (2000)
Effects of display frames of reference on change detection and spatial judgments
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 181

> In previous experiments, we showed that subjects' biases influenced how they examined a pictorial rendition of a virtual battlefield. Subjects could be immersed in the battlefield, having a view like that of a soldier standing on the ground, or they could have a wider view, in which they were elevated above the ground, observing the virtual battleground as one would from a helicopter. Subjects with the immersed view were less likely to include information outside their field of view in their answers to questions about the battlefield, even though a small map inset, on which the information appeared, was placed at the top of the immersed view. To explore these findings further, we showed subjects three virtual battlefields, one viewed from an elevated perspective and two from immersed perspectives. In one immersed perspective, subjects controlled panning; in the other, panning was controlled by automation. We

wondered if automatic panning would improve subjects' knowledge of information outside their initial field of view. We found that automatic panning did not improve subjects' knowledge; attention seemed to be captured by the initial field of view seen by subjects. We discuss how these findings are relevant to three-dimensional, computer-generated renditions of a battlefield.

Thomas, L. C., & Wickens, C. D. (2000)
Effects of display frames of reference on spatial judgments and change detection
Tech. Rep. No. ARL-00-14/FEDLAB-00-4, Savoy, Illinois: University of Illinois at Urbana-Champaign, Aviation Research Laboratory Institute of Aviation

We compared three types of computer-generated displays of battlefield information in order to address the possible influence of any of four potential causes of display-induced cognitive tunneling. The battlefield information was presented as realistic terrain imagery. Participants viewed the imagery from an elevated perspective, as would an observer in an airplane, or immersed within the terrain; if immersed, participants could view parts of the terrain outside the field of view by panning left or right or by allowing the apparatus to automatically pan the view. The participants' tasks were to make spatial judgments, provide counts of visible enemy units, detect changes in the units, and select paths through the environment. Participants in the auto-panning immersed group performed the tasks more poorly than those in the self-panning immersed group and the elevated perspective group. In addition, all groups exhibited cognitive tunneling by their tendency to note changes in centrally located information more accurately than peripheral information.

Uckun, S., Tuvi, S., Winterbottom, R., & Donohue, P. (1999)
OWL: A decision-analytic war-gaming tool
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 133–137

OWL is a decision-analytic wargamer that is used to evaluate the benefits and risks of multiple friendly courses of action. OWL uses stochastic simulation principles to evaluate alternate outcomes of a battle, given uncertainty in the information available about friendly forces, the enemy, mission, weather, and the terrain. OWL is designed as a postprocessor for Fox, a tool that evaluates thousands of potential courses of action and selects a small number of plausible ones.

Vassiliou, M. S. (2000)
The ARL Displays FedLab: A partnership between industry, government, and academia
Proceedings of the IEEE Aerospace Conference, 6, 521–529

> In order to better accomplish its mission of serving Army R&D needs, the U.S. Army Research Laboratory has pioneered the concept of FedLab. This is a new method of conducting government-sponsored research in which ARL is an active participant in and manager of a research consortium involving various industrial and academic partners. A FedLab is funded via a new instrument, the cooperative agreement. The industrial and academic laboratories effectively become virtual divisions of ARL, enhancing and complementing its internal capabilities. Three FedLab consortia were established in 1996 to perform research in areas related to the future "digitization of the battlefield." One of these is the Advanced Displays and Interactive Displays (ADID) consortium, a 5-year basic research effort to develop new technologies in human-computer interaction and related areas. The consortium, which is managed by a committee of representatives from all members, reports to a program manager at ARL. The management committee and ARL jointly prepare annual research plans and work to ensure the relevance of the research to customers in the Army. Great care is taken to see that resources are committed to the technology transition process. The ADID consortium brings together investigators with a unique mix of skills in computer science, engineering, and human factors, with orientations ranging from very fundamental research to highly practical and pragmatic work.

Vassiliou, M. S., & Huang T. S. (Eds.) (2001)
Computer-Science Handbook for Displays, Summary of Findings from the Army Research Lab's Advanced Displays & Interactive Displays Federated Laboratory
Adelphi, MD: Army Research Laboratory

> The purpose of this handbook is to distill and synthesize some of the salient points developed over the 5-year run of the U.S. Army Research Laboratory's Advanced Displays and Interactive Displays consortium. The purpose of the consortium was to perform fundamental research to develop new technologies in human-computer interaction, information display, decision analysis, and related areas to support the Army of the future. The consortium was concerned with the software, algorithms, and human factors science and engineering required for the effective display and presentation of information and knowledge about a broad variety of hardware platforms. This book, as its title implies, concentrates on the computer-science aspects of the research, while the companion volume concentrates on human factors.

Vassiliou, M. S., Sundareswaran, V., Chen, S., Behringer, R., Tam, C., Chan, M., Bangayan, P., & McGee, J. (2000)
Integrated multimodal human-computer interface and augmented reality for interactive display applications
Proceedings of SPIE's International Society for Optical Engineering, **4022**, 106–115

> We describe new systems for improved integration of multimodal human-computer interaction and augmented reality for a diverse array of applications in future advanced cockpits, tactical operations centers, and other settings. We have developed an integrated display system featuring capabilities such as (1) speech recognition of several, concurrent speakers via standard air-coupled microphones or novel throat-coupled sensors (developed at the U.S. Army Research Laboratory for increased noise immunity); (2) lip reading for improving speech recognition accuracy in noisy environments; (3) three-dimensional (3-D) spatialized audio for improved display of warnings, alerts, and other aural information; (4) wireless, coordinated hand-held PC control of a large display; (5) real-time display of data and inferences from wireless integrated networked sensors with on-board signal processing and discrimination; (6) gesture control with disambiguated point-and-speak capability; (7) head and eye tracking coupled with speech recognition for look-and-speak interaction; and (8) integrated tetherless augmented reality on a wearable computer. The various interaction modalities (speech recognition, 3-D audio, eye tracking, etc.) are implemented as "modality servers" in an Internet-based client-server architecture. Each modality server encapsulates and exploits commercial and research software packages, presenting a socket network interface that is abstracted to a high-level interface, minimizing both vendor dependencies and required changes on the client side as the server's technology improves.

Vassiliou, M. S., Sundareswaran, V., Chen, S., & Wang, K. (1999)
Multimodal HCI integration
Society of Automotive Engineers, 1999 World Aviation Congress, Report No. 99WAC-149

> A multipurpose test bed for integrating user interface and sensor technologies has been developed, based on a client-server architecture. Various interaction modalities (speech recognition, three-dimensional audio, pointing, wireless hand-held PC-based control and interaction, sensor interaction, etc.) are implemented as servers, encapsulating and exposing commercial and research software packages. The system allows users to interact with large and small displays via speech commands as well as by pointing, spatialized audio, and other modalities. Simultaneous and independent speech recognition for two users is supported; users may be equipped with conventional acoustic or new body-coupled microphones.

Walrath, J., Gurney, J., & Voss, C. (1997)
All in favor, say eye

*Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, (Pt. 1) 57–62*

> The U.S. Army Research Laboratory is developing a virtual environment
> for a battlefield visualization system (BVS), which currently requires that
> all input from the user be made through pull-down menus controlled by
> a mouse. This method does not allow the user to fully exploit the
> medium. One strategy for a less intrusive, more intuitive way of
> interacting with the BVS is to develop a natural language interface. One
> of the greatest technical challenges in developing such an interface is that
> individual sentences are frequently ambiguous. Often, this ambiguity
> results from the user's reference to one of a group of objects on the
> display (e.g., the helicopter) or a position (e.g., here, there, beside, etc.).
> Our approach to this challenge is to combine linguistic analyses of the
> user's speech with nonlinguistic (in this case, visual) information about
> which elements in the scene are most salient to the user at the time of the
> spoken request or command. We will track, in tandem, two distinct
> modes of user input: the user's speech and point of gaze in the visual
> scene. We will use an eye tracker to identify the objects at which the user
> is looking, thus overcoming the lower precision (ambiguity) in the user's
> vocal expressions. We hypothesize that if the user's utterance refers to an
> object or contains a spatial referent, then point-of-gaze data, collected just
> before or during the utterance, will provide reliable information about
> which object or location was intended.

Walrath, J., Voss, C., & Gurney, J. (1998)
Towards a hands-free interface: Tracking natural eye movements during discourse

*Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 27-32*

> Our work is part of a larger research effort to construct a hands-free
> interface for a virtual reality battlefield visualization system. The work
> was conducted to determine if eye gaze can supplement linguistic
> analyses in extracting meaning from natural language discourse. Subjects
> viewed a map on a computer monitor while an eye-movement system
> determined their points of gaze. The subject's task was to describe a route
> marked on the electronic map to a second participant (the cohort), who
> had an equivalent paper map but no marked route. The cohort's task was
> to draw the route on the paper map, based solely on verbal interaction
> with the subject. The subject's verbal discourse and eye movement
> information were analyzed. Results indicate that point-of-gaze data
> provide valuable nonlinguistic information about which elements in the
> visual scene are important to the user during speech production.

Wang, R., & Huang, T. (1999)
Fast camera motion analysis in MPEG domain

Proceedings 1999 International Conference on Image Processing, 3, 691–694

Camera motion estimation is crucial for video analysis and for object-tracking query systems if the motion of an object needs to be neutralized before the object can be analyzed. As the amount of video data contained in formats such as MPEG-1 and MPEG-2 grows, it increasingly makes more sense to perform motion estimation on the compressed data. Much work has gone into analyzing uncompressed video, but the time to uncompress and analyze data is simply too great for many large video databases. This paper presents a recursive, outlier-rejecting least square algorithm for parametric camera estimation for visual information in MPEG-1 and MPEG-2 formats. The algorithm has a very low time complexity, and experimental results show that it works much faster than real-time playback and consumes few system resources. Experiments on synthesized and real-world video clips show that the algorithm is effective. Experiments are also conducted on a large set of real-world video clips, and a query system is built in the process.

Wang, K, K, Sundareswaran, V. S., McGee, J. H., Chen, S. L., & Chan M. T. (2001)
An abstracted interface for human-computer interaction pointing devices

Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 117

An increasing number of human-computer interaction (HCI) pointing devices are becoming available to application developers and users. These include commercial off-the-shelf devices (e.g., mouse, tablet, wand) and more exotic devices developed within the advanced displays and interactive displays consortium (e.g., head motion-compensated eye tracker, laser pointer, gesture recognizer), which run on a variety of platforms. There is a need to integrate these various devices across multiple platforms (e.g., integrating an application running on a Silicon Graphics, Inc., [SGI] computer with tablet service running on a PC that controls the SGI application), which may be fulfilled through a versatile network interface. At the Rockwell Scientific Company (RSC), we are developing several pointing services (including head tracking, eye tracking, tablet), which are accessible through network interfaces. As an initial step toward a uniform network interface, we have developed a device-neutral Coordinate Space Transform (CST) server that abstracts specific pointing device data into a generic pointing device form. The CST server is designed to receive data from a pointing server, transform the data as appropriate for a given application, and serve the transformed data to the client application. We describe an RSC-developed look-and-speak application, wherein ultrasonic head tracking is used to effect head motion-compensated eye tracking, allowing the user to look across multiple physical displays. As the user turns his head or moves his eyes, an application tracks which display (if any) is currently in visual regard

and uses that display as the object for his speech commands. Through the use of speech, the user can request the display of any available media and can pause, seek, and resume playback of media. Gaze is used to resolve which of three available displays the user is referring to while issuing speech commands.

Weber, T. A., Kramer, A. F., & Karni, O. (1998)
Tracking visual attention with event-related brain potentials
Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 33–38

> There is a widely accepted view in the psychological literature that visual attention can be allocated only to a specific and limited area of the visual field at any one time. A number of techniques have also been developed, both psychophysical and physiological, to determine where in the visual field attention has been allocated. We used one of the physiological techniques known as steady state visually evoked potentials (SSVEPs) to determine non-invasively where attention has been directed in the visual field, regardless of eye position. This research is based on the finding that recorded SSVEPs to an attended location are larger in amplitude than the SSVEPs to an unattended location. The SSVEP technique involves recording the brain's electrical activity to irrelevant background flashes via electroencephalography (EEG) while the subject performs some visually oriented task, such as monitoring for the occurrence of a target. The irrelevant background flashes occurred in two locations on a video screen at two different frequencies of modulation, while the targets were monitored in only one of the pre-specified locations. We used the recorded EEG to construct SSVEP waveforms to determine the location of attention and then developed a metric for using the SSVEPs to show the allocation of attention to a specified location. Two essential aspects of this project are that (1) we used non-intrusive recording (i.e., sensors placed externally on the scalp), and (2) we used non-obtrusive recording (i.e., no overt response to the background flashes was necessary). We discuss potential applications of this technique to target detection, vigilance monitoring, and training. Long-term goals for this project include extending this technique from off-line analysis to real-time determination of a visually attended location and real-time assessment of vigilance monitoring.

Weber, T. A., Kramer, A. F., & Miller, G. A. (1997)
Selective processing of superimposed objects: An electrophysiological analysis of object-based attentional selection
Biological Psychology, 45(1-3), 159–142

> A study investigated whether object-based attentional selection occurs from grouped array or spatially invariant representations. Eighteen college students were presented with colored objects and asked to judge whether a particular color-shape conjunction was present, regardless of

whether the color and shape were part of a single object (same object condition) or occurred on two different objects (different object condition). Reaction times (RTs) and accuracies were recorded for subjects' judgments. Event-related brain potential components, particularly the P1 and N1, were elicited both from the presentation of the target objects and from a post-display probe that was used as an index of spatial attention. Consistent with predictions of object-based selection models, RTs and accuracies were faster in same object than in different object trials. N1s elicited by the target objects and P1s elicited by the post-display probes discriminated between same and different object trials when the two target objects were superimposed. These data are consistent with the proposal that object-based selection is spatially mediated, even for partially overlapping objects. The data are discussed in terms of space- and object-based models of visual selective attention.

Wesler, M. Mc., Darkow, D., & Marshak, W. (2000)
The effects of training on a multi-modal common metric for information-based displays
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 59–63

> We investigated the effects of training as an initial step in verifying the cognitive implications associated with multi-modal displays and their evaluation using a common metric (CM) algorithm. Performance improved over the repeated performance of the task. However, no significant difference was found between distributed and massed practice on a detection task. Two parameters used to compute the CM changed because of the cognitive effects of training. The anticipated change in performance at lower signal-to-noise ratios was observed, along with an unanticipated change in performance slope. The impact of cognition on computing the metric and the direction of future research into the measure are discussed.

Wickens, C. D. (2000)
Human factors in vector map design: The importance of task-display dependencies
Journal of Navigation, 53, 54–67

> The role of human factors in map design is to serve as a mediator between the technology availed by electronic digital maps (particularly vector maps) and the many tasks performed by the user. Simply put, no single map is best suited for all tasks. The appropriate relationship between map and task is, in turn, mediated by a series of information processing principles, articulated by the engineering psychologist. The field of engineering psychology is on the threshold of providing computational models, based on these principles, which will supply guidance to the map designer in regard to the circumstances that make one map format better than another for a particular task. This paper

114

describes these principles as applied to two domains of vector map design: the domain of three-dimensional maps and the domain of database overlay.

Wickens, C. D. (2000)
The when and how of using 2-D and 3-D displays for operational tasks
Proceedings of the Human Factors & Ergonomics Society 44th Annual Meeting, 3 , 403–406

>Three different canonical viewpoints into a three-dimensional (3-D) domain are defined to create a taxonomy of 3-D displays. I then show how the information processing demands of each display viewpoint provide benefits and imposes costs on four categories of tasks, involving travel, image matching or situation awareness, visual search, and precise judgments. These task-display interactions are illustrated from experiments in aviation display design, battlefield judgments, and data visualization. Conclusions are offered regarding two possible ways of addressing the task-display interactions in design.

Wickens, C. D. (2001)
Situation awareness on the battlefield: An integration of FedLab research
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & *Interactive Displays Consortium, 65–70*

>I review and synthesize FedLab research products that address how soldiers monitor and integrate multiple sources of dynamic information to attain an accurate assessment of the battlefield situation. Substantial problems are encountered in noticing changes in battlefield information, but computer automation can support change detection. Cognitive biases in integrating semi-reliable information, over space and over time, are then discussed. Biases in cognitive effort conservation and over-relying on automation appear to be more serious than biases in over-weighting salient information.

Wickens, C. D., Kroft, P., Yeh, M. (2000)
Database overlay in electronic map design: Testing a computational model
Proceedings of the Human Factors & Ergonomics Society 44th Annual Meeting, 3 , 451–454

>In two experiments, participants answered questions about two geographical-spatial databases, which were displayed in different formats and at different levels of clutter. One experiment examined aviation information (traffic, weather, terrain), and the other examined information pertaining to a soldier's battlefield (troops, roads, rivers, and terrain). Databases were presented in five contrasting formats: overlay, spatially separated at small resolution, spatially separated at large resolution, highlighted, and with an interactive decluttering mode. Performance was evaluated in the context of the different information-processing mechanisms that were challenged or supported by the

different formats. The data revealed a linear effect of clutter on reaction time, a general benefit for highlighting, and a cost for interactive displays.

Wickens, C. D., Pringle, H. L., Merlo, J. (1999)
Integration of information sources of varying weights: The effect of display features and attention cuing

Tech. Rep. No. ARL-99-2/FEDLAB-99-1, Savoy, Illinois: University of Illinois at Urbana-Champaign, Aviation Research Laboratory Institute of Aviation

> This report reviews research in which multiple sources of variable reliability information are integrated for making diagnostic judgments or allocating resources. A framework for considering these experiments is provided, and some evidence is presented regarding the extent to which humans are "calibrated " in allocating processing proportionately to the ideal weights (i.e., reliability or importance) of information channels. Two generic sources of bias are identified. Attentional biases occur when more processing is given to less important channels at the expense of more important ones (i.e., a failure to allocate attention optimally). Trust biases occur when less than fully reliable information is offered more processing than is warranted (i.e., "overtrust"). The report also reviews and integrates the conclusion from a smaller number of specific studies that examined how multisource information processing is modulated by properties of the display of those sources. Two sources of display information are considered: attentional guidance (e.g., cuing), which directs attention to certain regions of the display, and reliability guidance, which explicitly displays the level of reliability of the information sources. Each type of information can induce the appropriate behavior from the user, either explicitly (e.g., by highlighting the important feature) or implicitly (by placing the important feature in the center of the display). Generalizations regarding the effectiveness of these display features are sought from the studies reviewed.

Wickens, C. D., & Rose, P. N. (2001)
Human Factors Handbook for Displays, Summary of Findings from the Army Research Lab's Advanced Displays & Interactive Displays Federated Laboratory

Adelphi, MD: Army Research Laboratory

> The purpose of this handbook is to distill and synthesize some of the salient points developed over the 5-year run of the U.S. Army Research Laboratory's Advanced Displays and Interactive Displays consortium. The purpose of the consortium was to perform fundamental research to develop new technologies in human-computer interaction, information display, decision analysis, and related areas to support the Army of the future. The consortium was concerned with the software, algorithms, and human factors science and engineering required for the effective display and presentation of information and knowledge on a broad variety of hardware platforms. This book, as its title implies, concentrates on the

human factors aspects of the research, while the companion volume concentrates on computer science.

Wickens, C. D., Thomas, L., Merlo, J., & Hah, S. (1999)
Immersion and battlefield visualization: Does it influence cognitive tunneling?
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 111–115

> Thirty officers at the U.S. Military Academy participated in a study in which a three-dimensional (3-D) exocentric display was compared with a 3-D immersed display, as a means of supporting situation awareness regarding an evolving battlefield scenario. The immersed viewpoint allowed 360° panning and was coupled with a small plan-view inset. A series of questions was asked about successive scenes as the movement to contact progressed. Results revealed that users of the immersed display demonstrated a form of "cognitive tunneling" in which they were overly influenced by information in the initially presented forward view, failing to adequately pan views behind them. The data speak to the advantage of 3-D exocentric displays.

Wickens, C. D., Thomas, L. C., & Young, R. (2000)
Frames of reference for the display of battlefield information: Judgment-display dependencies
Human Factors, 42, 660–675

> In two experiments, U.S. Army soldiers viewed computer-generated displays that presented battlefield information from three different frames of reference: a two-dimensional (2-D) plan view display (with contour lines), a three-dimensional (3-D) exocentric perspective display, and an interactive 3-D immersed display. In Experiment 1, soldiers made geographical judgments. The results suggested that both 3-D displays suffered from ambiguity of distance estimates but that the 3-D immersed display was most accurate for judging whether one location is directly visible from another. In Experiment 2, the 3-D exocentric display was compared with a 3-D immersed view, which included a small 2-D inset map, in a more continuous battlefield scenario in which judgments of enemy activity were made. The findings of 3-D ambiguity were replicated from Experiment 1. The accuracy of judgments of enemy activity suffered with the immersed display when information necessary to answer correctly did not appear in the initial forward view and required panning to acquire, thus reflecting the cognitive demands of integration across different views. This display also hindered soldiers' ability to report changes in enemy activity from one view to the next. The results of this research will help to provide guidelines for the appropriate choice of computer display technology to assist in designing battlefield visualization aids. Caution should be exercised in choosing immersive viewpoints.

Wickens, C. D., & Yeh, M. (1997)
Attentional filtering and decluttering techniques in battlefield map interpretation
Proceedings of the 1st Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, (Pt. 2), 35–42

> We compared the efficacy of color coding, intensity coding, and decluttering techniques for filtering information on battlefield maps. Eighteen subjects viewed computer-generated maps that contained five classes of information: roads, rivers, terrain, troops and fixed unit locations. Subjects answered a variety of questions based on this information. Classes of information were differentiated by color, by intensity coding, or in the decluttering condition, by displaying relevant information and removing other information by a single key press. The results revealed an advantage for the color coding over the control condition, with the intensity coding condition falling in between. The decluttering option was not helpful, since the cost of switching between present and absent information and of deciding whether needed information was present outweighed any advantages of computer filtering. Thus, human attentional filters appeared to be superior to computer filters with the electronic maps used. The implications of the results are discussed.

Wickens, C., & Yeh, M. (1997)
A comparison of emphasis techniques in electronic map displays: Attentional filtering vs. decluttering
Proceedings of the 41st Annual meeting of the Human Factors and Ergonomics Meeting, **2**, 1396

> Color coding, intensity coding, and decluttering were compared in order to determine their potential benefits for accessing information from electronic map displays. Eighteen subjects viewed electronic battlefield maps containing five classes of information discriminable by color, intensity, or in the decluttering condition, displayed or removed entirely by a key press. Subjects were asked questions requiring them to focus on objects within a class, integrate information across two commonly coded feature classes, or divide their attention between objects in different classes. The results suggested that the benefits of color and intensity coding appear to be in segregating the visual field rather than calling attention to the objects presented at a certain color or intensity. The decluttering option proved to be a comparative disadvantage; the decision time necessary to determine whether the information needed was present inflated the response time and outweighed the benefits of presenting less information on the display.

118

Wilkins, D. C., Mengshoel, O, J., Chernyshenko, O., Jones, P. M., Hayes, C. C., Bargar, R. (1999)
Collaborative decision making and intelligent reasoning in Judge Advisor Systems

Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, 1–9

> This paper examines the Raven and CoRaven decision-making tools, which are used to filter, interpret, and visualize large amounts of uncertain data. Raven and CoRaven are multimodal advisory decision aids that base their inferential reasoning on Bayesian networks. Human decision makers and information sources interact with these decision-making systems in many ways during their design, construction, refinement, and use. The collaborative aspects of using Raven and CoRaven are analyzed with the judge-advisor system model.

Wolfson, O. (Ed.) (1997)
Data management issues in mobile computing [Special section 1]

Mobile Networks and Applications, **2**

> This special section contains four articles that address some of the most important issues in adapting databases to a mobile computing environment.

Wolfson, O., Chamberlain, S., Dao, S., & Jiang, L. (1997, October)
Location management in moving objects databases

paper presented at the Second International Workshop on Satellite-Based Information Services, Budapest, Hungary

> The authors first introduce moving objects databases and their related research problems; they then concentrate on a particular problem, namely, reducing the information cost associated with a trip taken by a moving vehicle. The information cost of a trip consists of the overhead of position update messages, average uncertainty, and the deviation of the database position from the actual position of the object. Three position update policies are introduced: immediate linear policy (ILP), plain dead reckoning (PDR), and adaptive dead reckoning (ADR). ADR is shown to have a lower information cost than PDR.

Wolfson, O., & Huang, Y. (1998)
Competitive analysis of caching in distributed databases

IEEE Transactions on Parallel and Distributed Systems, **9**, 391–409

> The contributions of two models to distributed databases are described. The first is a model for evaluating the performance of data allocation and replication algorithms in distributed databases. The model is comprehensive in the sense that it accounts for I/O and communication costs, and because of reliability considerations, it accounts for limits on the minimum number of copies of the object. The model captures existing replica-management algorithms, such as read-one-write-all, quorum-consensus, etc. These algorithms are static in the sense that in the absence of failures, the copies of each object are allocated to a fixed set of

processors. The second model is concerned with the fact that in modern distributed databases (particularly in mobile computing environments), processors dynamically store and relinquish objects in their local database. An algorithm is introduced for automatic dynamic allocation of replicas to processors. Using the new model, the authors compare the performance of the traditional read-one-write-all static allocation algorithm with the performance of the dynamic allocation algorithm. The relationship between the communication cost and I/O cost for static allocation is superior to that for dynamic allocation.

Wolfson, O., Lelescu, A., & Xu, B. (1999, September)
Retrieval of collaborative work from multimedia databases using relevance feedback

paper presented at the Proceedings of the Symposium on String Processing and Information Retrieval, Cancun, Mexico

In this paper, we address the problem of retrieving stored multimedia presentations by using relevance feedback. We model multimedia presentations using a crisp relational or object-oriented database, augmented with a text attribute. We also introduce a language for retrieval by content from such databases. The language is based on fuzzy logic. We also introduce a method for query refinement that uses relevance feedback provided by the user.

Wolfson, O., Sistla, P., Xu, B., Zhou, J., Chamberlain, S., Yesha, Y., & Rishe, N. (1999)
Tracking moving objects using database technology in DOMINO

Lecture Notes in Computer Science, **1649,** 112–120

Methods are discussed for overcoming the limitations of computerized database management systems (DBMSs) when they contain information about moving ground or air vehicles. DBMSs have problems managing large amounts of continuously changing data (e.g., changes in the location of a large number of vehicles), representing spatial data (e.g., vehicles near a common destination), and handling imprecise information about a vehicle's location. The authors discuss how their Database fOr MovINg Objects (DOMINO) project will resolve these issues.

Wolfson, O., Xu, B., Chamberlain, S., & Jiang, L. (1998)
C hallenges and approaches in motion databases

Proceedings of the 14th International Conference on Advanced Science and Technology, 182–194

Abstract not available

Wolfson, O., Xu, B., Chamberlain, S., & Jiang, L. (1998)
Moving objects databases: Issues and solutions

Proceedings of the 10th International Conference on Scientific and Statistical Database Management, 111–122

The authors report about research into the tracking of moving objects and their locations in a database, such as the location of moving taxicabs in a

city. Currently, moving objects database applications are being developed in an *ad hoc* fashion. Database management system (DBMS) technology provides a potential foundation upon which to develop these applications; however, DBMSs are currently not used for this purpose because a critical set of capabilities needed by moving objects database applications is lacking in existing DBMSs. The objective of the current project, called DOMINO (databases for moving objects), is to build an envelope containing these capabilities on top of existing DBMSs. Problems and proposed solutions are discussed.

Wright, S. (1999)
Effects of computer displayed color characteristics on individuals
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 171

Fifty participants subjectively rated 25 five-color samples for pleasantness, arousal, and dominance. The color samples were based on combinations of five different hues (blue, green, red, yellow, and purple), three saturation levels (low, medium, and high), and three brightness levels (low, medium, and high). These combinations were varied in a methodical manner along a predetermined scale. The colors were specified by RGB (red, green, blue) values, HSV (hue, saturation, value) values, and Munsell notation. Based on results from these ratings, numeric models were developed through regression analysis to predict the pleasantness and arousal levels of screen background colors, based on the color's characteristics. These models may be used to determine choice of background and foreground colors for information displays that require the user to experience a predetermined level of arousal.

Wright, S. (1999)
The impact of color characteristics on visual search pattern
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 173

The primary purpose of this research was to determine whether visual search patterns were affected by different color combinations of brightness and saturation. A secondary purpose was to determine whether individuals usually start scanning graphic information in the same position. Colors with high levels of brightness and saturation were expected to draw the eye, thus changing the visual search pattern. An eye scanner was used to examine the search patterns of 15 subjects while they scanned an array of 16 variously colored icons for a previously designated icon. Results show that eye-scanning patterns do change, based on the color combinations of surrounding icons. Results from this experiment should influence the color characteristics of icons and symbols requiring immediate attention on a display.

Wu, Y, & Huang, T. (1999)
Capturing articulated human hand motion: A divide-and-conquer approach

Proceedings of the Seventh IEEE International Conference on Computer Vision, 1,
606–611

The use of the human hand as a natural computer interface device has inspired research in the modeling, analyzing, and capturing of the motion of the articulated hand. Model-based hand motion capturing can be formulated as a large nonlinear programming problem, but this approach is plagued by local minima. An alternative is to use analysis by synthesis in searching a huge space, but the results are inexact and the computation expensive. In this paper, articulated hand motion and finger motion are decoupled, and a new two-step iterative model-based algorithm is proposed to capture articulated human hand motion. A proof of convergence of this iterative algorithm is given. In our proposed work, the decoupled global hand motion and local finger motion are parameterized by the three-dimensional (3-D) hand pose and the state of the hand. Hand pose determination is formulated as a least median of squares problem rather than the non-robust least squares (LS) problem, so that 3-D hand pose can be reliably calculated even if there are outliers. Local finger motion is formulated as an inverse kinematics problem. A genetic algorithm-based method is proposed as an effective method of finding a sub-optimal solution to the inverse kinematics problem. Our algorithm and the LS-based algorithm are compared in several experiments. Both algorithms converge when local finger motion between consecutive frames is small. When large finger motion is present, the LS-based method fails, but our algorithm can still successfully estimate the global and local finger motion.

Wu, Y., & Huang, T. (1999)
Human hand modeling, analysis and animation in the context of HCI

Proceedings 1999 International Conference on Image Processing, 3, 6–10

The use of the human hand as a natural interface device serves as a motivating force for research in visual analysis of highly articulated hand movement. Since hand motion covers a huge domain, the scope of this paper is limited to the developments of three-dimensional (3-D) model-based approaches. Numerous 3-D models that have been used to analyze hand motion are studied. Various approaches to articulated motion analysis are discussed. Some realistic synthesis methods are also included in this paper. We conclude with some thoughts about future research directions.

Wu, Y., Liu, Q., & Huang, T. S. (2000)
Tracking, analyzing and recognizing gesture commands
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays &
Interactive Displays Consortium, 183

> Conventional input devices, such as keyboards, mice, wands, and joysticks, are not natural and convenient for many current interactive applications such as virtual environments and robot control. Hand gestures can serve as a more natural way for humans to interact with machines. There are several challenges in implementing hand gesture recognition, including robust hand localization and tracking, hand posture recognition, and temporal gesture recognition. To track a hand, our system extracts changes in the color distribution of the hand as it moves through space. We combine principal component analysis and multiple discriminate analysis to extract some of the discriminating features resulting from hand motion. Our approach to temporal gesture commands recognition is based on dynamic self-organizing map that has been successfully used in speech and handwriting recognition. Further research is needed to integrate these parts and optimize the system.

Wu, J. J., Sharma, R., & Huang, T. S. (1998)
Analysis of uncertainty bounds due to quantization for 3-D position estimation using multiple cameras
Optical Engineering, **37,** 280–292

> An important source of error when one is estimating the three-dimensional position of a point from two (stereo), three (trinocular), or more cameras is quantization error on the image planes. In this paper, we are concerned with bounding the quantization errors when using multiple cameras defined in terms of uncertainty regions in three dimensions. We use a geometric error analysis method that models the quantization error as projected pyramids and the uncertainty region as an ellipsoid around the polyhedron intersection of the pyramids. We present a computational technique for determining the uncertainty ellipsoid for an arbitrary number of cameras. A numerical measure of uncertainty bond, such as the volume of the ellipsoid, can then be computed for aiding camera placement, trajectory planning, and various other multiple camera applications.

Yeh, M., Brandenberg, D., & Wickens, C. W. (2000)
Up or down? A comparison of helmet-mounted display and hand-held display tasks with high clutter imagery
Tech. Rep. No. ARL-00-11/FED-LAB-00-3, University of Illinois at Urbana-Champaign, Aviation Research Lab Institute of Aviation

> We examined the trade-offs between the costs of increased clutter that result by overlaying complex information onto the forward field of view using a helmet-mounted display (HMD) with the cost of scanning when presenting this information on a hand-held display (HHD). Eight

National Guard personnel were asked to detect, identify, and give azimuth information for targets hidden in terrain presented in a simulated far domain environment while they performed a monitoring task in the near domain using either an HMD or HHD. The results revealed that the costs of clutter outweighed the cost of scanning as the amount of information that needed to be inspected increased. The presentation of cuing, which guided attention to a large region around the target, facilitated detection without imposing the costs of attentional tunneling.

Yeh, M., & Wickens, C. D. (1999)
Visual search and target cuing with augmented reality: A comparison of head-mounted with hand-held displays
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 105–109

We conducted a study to determine the effects of target cuing and conformality with a hand-held display (HHD) or helmet-mounted display (HMD) on visual search tasks requiring focused and divided attention. Eleven military subjects were asked to detect, identify, and give azimuth information for targets hidden in terrain presented in a simulated far domain environment, while they concurrently monitored a nearby domain using either an HHD or HMD. The results showed that the presence of cuing aided the target detection task for expected targets but drew attention away from unexpected targets in the environment. This effect was reduced when subjects used the HHD. Additionally, the results showed that the presence of cuing hindered performance on the secondary task.

Yeh, M., & Wickens, C. D. (2000)
Attentional and trust biases in augmented reality: Examining the trade-offs of interactivity, image realism, and the presentation of cuing symbology
Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 53–57

This experiment seeks to examine the effects of three variables (cue reliability, image reality, and user's interactivity) on user's attention to and trust in target cuing. Sixteen military personnel were asked to detect targets camouflaged in scenes, presented at two levels of image reality. To aid them in target detection, cuing was presented for some of the targets; the reliability of the cuing information was manipulated at two levels (100% and 75%). Half of the subjects actively navigated through the terrain; the other half passively viewed the passing terrain as their course was guided by an autopilot. The results showed that the presence of cuing aided target detection for expected targets but drew attention away from the presence of unexpected targets. This attentional tunneling was mediated by cue reliability: unexpected targets presented in conjunction with a cued target were detected more often when cuing was only

partially reliable. Neither image reality nor interactivity directly influenced trust in the cuing. Instead, the effect of enhanced reality was attributable to the lower visibility of the target in the scene, and the influence of interactivity was attributable to increased resource demand, which modulated performance in the presentation of unreliable cuing.

Yeh, M., & Wickens, C. D. (2001)
Attentional filtering techniques in the design of battlefield maps: Examining the use of color, intensity coding, and decluttering
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 35–40

> In a series of experiments, color coding, intensity coding, and decluttering were compared in order to assess their potential benefits for emphasizing critical information on electronic map displays. Participants viewed electronic battlefield maps containing five classes of information discriminable by color or intensity or in the decluttering condition, displayed or removed entirely by a key press. Participants were asked questions requiring them to focus on objects within a class (i.e., objects presented at the same color or intensity) or to divide their attention between objects in different classes (i.e., objects presented at different colors and intensities). The results suggested that the benefits of color and intensity coding appear to be in segregating the visual field rather than in calling attention to the objects presented at a certain color or intensity. The cost of decluttering outweighed the benefits of presenting less information on the display or even allowing map users to customize their displays.

Yeh, M., Wickens, C. D. Merlo, J. L., & Brandernburg, D. L. (2001)
Examining the clutter-scan trade-off with high clutter imagery: A comparison of helmet-mounted versus hand-held display presentation
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 29–33

> The current experiment was designed to examine the trade-offs between the costs of increasing clutter by overlaying complex information onto the forward field of view via a helmet-mounted display (HMD) versus the cost of scanning when this information is presented on a hand-held display (HHD). Eight National Guard personnel were asked to detect, identify, and give azimuth information for targets hidden in terrain presented in a simulated far domain environment while they performed a monitoring task in the near domain using either an HMD or HHD. The results revealed that the costs of clutter outweighed the cost of scanning in the presentation of complex visual information.

Yeh, M., Wickens, C. D, & Seagull, F. J. (1998)

Effects of frame of reference and viewing conditions on attentional issues with helmet-mounted displays

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 107–114

> We conducted a study to examine the issues of frame of reference, target expectancy, target cuing, and viewing condition (i.e., one eye versus two) in the design of helmet-mounted displays (HMDs) in order to determine their effects on focused and divided attention tasks. Sixteen subjects (8 civilian, 8 military) were asked to detect, identify, and give location information for targets hidden in images of terrain presented in the far domain while they performed a monitoring task in the near domain. The results showed that the presence of cuing aided target detection for expected targets but drew attention away from the presence of unexpected targets in the environment. Attentional capture was mediated by frame of reference: unexpected targets were detected more often when subjects searched with the HMD possessing conformal imagery than when the imagery was not conformal. Viewing a display with both eyes produced a very slight benefit in target detection.

Yeh, M., Wickens, C. D., & Seagull, F. J. (1999)

Conformality and target cuing: Presentation of symbology in augmented reality

Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society, **2,** 1526–1530

> We conducted a study examining several issues in the design of see-through helmet-mounted displays to determine their effects on tasks of focused and divided attention. These issues are frame of reference (world referenced versus screen referenced), target expectancy, target cuing, and viewing condition (i.e., one eye versus two). Sixteen subjects (8 civilian, 8 military) were asked to detect, identify, and give azimuth information for targets hidden in terrain presented in the far domain (i.e., the world) while they performed a monitoring task in the near domain (i.e., the display). The results showed that the presence of cuing aided target detection for expected targets but drew attention away from unexpected targets in the environment. However, analyses support the observation that this effect can be mitigated by the use of world-referenced symbology. Displaying symbology to two eyes provided a slight benefit for target detection when the target was cued.

Yeh, M., Wickens, C.D., & Seagull, F.J. (1999)

Target cuing in visual search: The effects of conformality and display location on the allocation of visual attention

Human Factors, **41,** 524–542

> Two experiments were performed to examine how frame of reference (world referenced versus screen referenced) and target expectancy can

126

modulate the effects of target cuing in directing attention for see-through helmet-mounted displays (HMDs). In the first experiment, the degree of world referencing was varied by the spatial accuracy of the cue; in the second, the degree of world referencing was varied more radically between a world-referenced HMD and a hand-held display (HHD). Participants were asked to detect, identify, and give azimuth information for targets hidden in terrain presented in the far domain (i.e., the world) while they performed a monitoring task in the near domain (i.e., the display). The results of both experiments revealed a cost-benefit trade-off for cuing so that the presence of cuing aided the target detection task for expected targets but drew attention away from the presence of unexpected targets in the environment. Analyses support the observation that this effect can be mediated by the display: the world-referenced display reduced the cost of cognitive tunneling relative to the screen-referenced display in Experiment 1; this cost was further reduced in Experiment 2 when participants were using an HHD. Potential applications of this research include important design guidelines and specifications for automated target recognition systems as well as any terrain-and-targeting display system in which superimposed symbology is included, specifically in assessing the costs and benefits of attentional cuing and the means by which this information is displayed.

Yu, H., Mehrotra, S., Winkler, R., Ho, S. S., Gregory, T. C., & Allen, S. D. (1999)
Integration of SATURN system and VGIS
Proceedings of the 3rd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 59–63

The Spatiotemporal Uncertainty Reasoning (SATURN) system, currently being developed, is being integrated with the Virtual Geographic Information System (VGIS) system in an effort to improve VGIS performance and scalability to complex dynamic environments, as well as to enhance its functionality as a collaborative planning tool. We added three new components to VGIS: a spatiotemporal object manager, a performance monitor, and a task database. The spatiotemporal object manager uses SATURN techniques for indexing dynamic multidimensional (spatiotemporal) objects to support effective and efficient object traversal during visualization. The performance monitor adjusts the resource allocation between VGIS components and adaptively adjusts image quality to guarantee bounded visualization performance. The task database extends VGIS as a tool for collaborative planning. Performance results illustrate that the SATURN techniques for object management and the performance monitor significantly improve VGIS performance, allowing it to scale to complex scenarios with a large number of dynamic objects.

Zahorik, P., Tam, C., Wang, P., Bangayan, P., & Sundareswaran, V. (2001)
Localization accuracy in 3-D sound displays: The role of visual feedback training
Proceedings of the 5th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 17–22

> Using an inexpensive headphone-based three-dimensional (3-D) display, six listeners localized sound before, during, and after perceptual feedback training. The training paired auditory and visual feedback with the position of the correct sound source. We show that feedback training markedly improved localization accuracy, with the largest improvements resulting from listener's enhanced abilities to distinguish sources in front from sources behind. Further, these improvements were not transient short-term effects but appear to last a number of days between training and testing sessions. These results suggest that simple and relatively short periods of perceptual training can effectively mitigate technical deficiencies in low-cost 3-D sound systems because non-individualized head-related transfer functions are used.

Zeller, M., Phillips, J. C., Dalke, A., Humphrey, W., Schulten K., Sharma, R., Huang, T. S., Pavlovic, V. I., Zhao, Y., Lo, Z., & Chu, S. (1997)
A visual computing environment for very large scale biomolecular modeling
IEEE International Conference on Application-Specific Systems, Architectures and Processors, 3–12

> Knowledge of the complex molecular structures of living cells is being accumulated at a tremendous rate. Key technologies enabling this success have been high-performance computing and powerful molecular graphics applications; however, the technology is beginning to lag in the face of challenges posed by the size and number of new structures and by the emerging opportunities in drug design and genetic engineering. For interactive modeling of biopolymers, a visual computing environment is being developed that links a three-dimensional (3-D) molecular graphics program with an efficient molecular dynamics simulation program executed on remote high-performance parallel computers. The system will be ideally suited for distributed computing environments because it uses both local 3-D graphics facilities and the peak capacity of high-performance computers for interactive biomolecular modeling. For creating an interactive 3-D environment, various input methods are possible. Three are explored: (1) a six-degree-of-freedom "mouse" for controlling the space shared by the model and the user, (2) voice commands monitored through a microphone and recognized by a speech recognition interface, and (3) hand gestures, detected through cameras and interpreted with computer vision techniques. Controlling 3-D graphics connected to real-time simulations and using voice with suitable language semantics, as well as hand gestures, promise great benefits for many types of problem-solving environments. Our focus on structural

biology takes advantage of existing sophisticated software, provides concrete objectives, defines a well-posed domain of tasks, and offers a well-developed vocabulary for spoken communication.

Zeller, M., Schulten, K., & Sharma, R. (1997)
Learning the perceptual control manifold for sensor-based robot path planning

Proceeding of the IEEE International Symposium on Computational Intelligence in Robotics and Automation, 48–53

> The perceptual control manifold is a concept that extends the notion of the robot configuration space to include sensor feedback for robot motion planning. In this paper, we propose a framework for sensor-based robot motion planning which uses the topology-representing network algorithm to develop a learned representation of the perceptual control manifold. The topology-preserving features of the neural network lend themselves to yield (after learning) a diffusion-based path-planning strategy for flexible obstacle avoidance. Simulations of path control and flexible obstacle avoidance demonstrate the feasibility of this approach for motion planning and illustrate the potential for further robotic applications.

Zhang, B., & Huang, T. (2000)
Evaluation of a hidden Markov model-based audio-visual speech recognizer on NATO-RSG-10 noise database

Proceedings of the 4th Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 83–85

> The performance of acoustic speech recognizers (ASRs) degrades significantly when there are mismatches between the training and operating conditions. Using signal processing techniques to reduce such mismatches has been the main research approach for enhancing ASR performance. Not until recently has the technique of using visual speech information to reduce the mismatch been investigated vigorously. The major issues for successfully incorporating visual cues into an ASR are first, acquiring the visual speech signal reliably and efficiently; second, integrating the visual and acoustic cues in an optimal way so that the mismatches in each modality will be maximally compensated for by the presence of the other modality; third, creating a much-needed database for analyzing audio-visual speech. This paper focuses on a hidden Markov model-based data fusion architecture. The temporal synchronous/asynchronous effects of the audio-visual speech sequences are learned and built into the speech models via a maximum likelihood estimation procedure. The bimodal ASR is evaluated by a speaker-independent connected digit recognition experiment under noise contamination from the NATO-RSG-10 noise database. The ASR demonstrates consistent recognition accuracy improvement over any single modal ASR for a wide range of signal-to-noise ratios.

Zhuang, Y., Huang, T. S., & Mehrotra, S. (1998)
A conceptual framework for multimedia reasoning

Proceedings of the 2nd Annual Federated Laboratory Symposium, Advanced Displays & Interactive Displays Consortium, 6–10

Artificial intelligence (AI) has been dominated by the physical symbolic system, in which symbolic information is used as the medium for reasoning. With this approach, information such as images, graphics, and video is transformed into symbols, which are fed into the AI system, and the symbolic result is transformed into its original media form. In this paper, we propose a new reasoning method called multimedia reasoning (MR), which is based on media such as text, image, video, audio, and so forth. We introduce the concept of multimedia transformation theory as a conceptual framework for multimedia reasoning. We discuss the importance and potential of MR in military applications.

Zhuang, Y., Rui, Y., Huang, T. S., & Mehrotra, S. (1998)
Adaptive key frame extraction using unsupervised clustering

Proceedings of the IEEE International Conference on Image Processing, 1, 866–870

Key frame extraction has been recognized as an important research issue in video information retrieval. Although progress has been made in key frame extraction, the existing approaches are computationally expensive or ineffective in capturing salient visual content. We first discuss the importance of key frame selection and then review and evaluate the existing approaches. To overcome the shortcomings of the existing approaches, we introduce a new algorithm for key frame extraction based on unsupervised clustering. The proposed algorithm is computationally simple and able to adapt to the visual content. The efficiency and effectiveness are validated by a large number of real-world videos.

Zhuang, Y., Rui, Y., Huang, T. S., & Mehrotra, S. (1998)
Applying semantic association to support content-based video retrieval

Fifth Very Low Bit-Rate Video Workshop (pp 45–48). University of Illinois, Urbana-Champaign

In the traditional approach to video retrieval, queries base their search on textual information (titles and keywords) annotated to the video. Since automated annotation is not yet available, generating keyword descriptors requires a great amount of labor and has proved to be unrealistic in applications. An approach that seems to be at the other extreme is using the low-level video content, such as color, texture, shape, and motion features, in an attempt to eliminate the necessity of keyword annotation. A preferable query form should include both keywords and video content. In this paper, we explore the semantic aspect based on video table of contents structuring. Closed captioning is used to extract a basic keyword set. Word-Net, an electronic lexical system, is used to provide semantic association. The approach has shown that retrieval performance is greatly improved.

Acronyms

ABATIS	Advanced Battlefield Architecture for Tactical Information Selection
ACAD	Alternative Courses of Action Display
ACD	associative configural display
ADAPT	(unknown)
ADID	Advanced Displays and Interactive Displays
ADR	adaptive dead reckoning
AHAS	(unknown)
AI	artificial intelligence
API	application programming interface
AR	augmented reality
AREAS	augmented reality system for evaluating assembly sequences
ARL	Army Research Laboratory
ASR	acoustic speech recognizer
AUI	adaptive user interface
BB	building block
BBN	Bayesian belief network
BFOS	Breiman, Friedman, Olshen, and Stone
BOA	Bayesian optimization algorithm
BRS	battlefield reasoning system
BVS	battlefield visualization system
C2V	command and control vehicle
CAD	computer-aided design
CBIR	content-based image retrieval
CHMM	coupled HMM
CIP	combat information processor
COA	course of action
COTS	commercial off-the-shelf
CST	coordinate space transform
DAISY	Design Aid for Intelligent Support Systems
DBMS	database management system
DEM	digital elevation map
DIVL	digital image/video library
DOMINO	database for moving objects
DSD	decision support display
DSS	decision support system
EA	evolutionary algorithm

ECOA	enemy COA
EBL	explanation-based learning
EEG	electroencephalography
FCOA	friendly COA
FORCES	Force Operational Readiness Combat Effectiveness Simulation
FP	fixed priority
FTL	future temporal logic
GA	genetic algorithm
GCMRD	gaze-contingent multi-resolution display
GDR	global dimensionality reduction
GEA	genetic evolutionary algorithm
GEC	genetic and evolutionary computation
GIS	geographic information systems
GiST	generalized search tree
GPS	global positioning system
GUI	graphical user interface
HCI	human-computer interaction
HCII	human-computer intelligent interaction
HES	hostile environment simulator
HHD	hand-held display
HMD	head-mounted display
HMM	hidden Markov model
HPC	hand-held personal computer
HTML	hypertext markup language
IGUANA	Intelligent Guidance and User-Adapted Interaction Agent
ILP	immediate linear policy
ISCAN	(unknown)
JAS	judge-advisor system
LDR	local dimensionality reduction
LINC	linkage identification by a nonlinearity check
LINC-AN	LINC-allowable nonlinearity
LIND	linkage identification by non-monotonicity detection
LS	least squares
MAP	maximum *a posteriori*
MARS	multimedia analysis and retrieval system
METD	minimum error tree decomposition
MHMM	multimodal HMM
MIR	multimedia information retrieval
MMR	(unknown)

MOST	moving objects spatio-temporal
MPEG	(unknown)
MR	multimedia reasoning
MRA	multi-resolution aggregate (tree)
MRF	Markov random field
MRWD	morphological representation of wavelet data
NC A&T	North Carolina Agricultural & Technical (State University)
OWL	(unknown)
PCD	process-centered display
PCM	perceptual control manifold
PDR	plain dead reckoning
QSR	qualitative spatial representation
R&D	research and development
RFR	relative force ratio
RMI	remote method invocation
RSC	Rockwell Scientific Company
RSS	rescaled signal subspace
RT	reaction time
SA	situational awareness
SATURN	spatiotemporal uncertainty reasoning
SGI	Silicon Graphics, Inc.
SNR	signal-to-noise ratio
SOA	stimulus onset asynchrony
SPIN	sensing positioning integrated network
SQL	structured query language
SSR	signal subspace rotation
SSVEP	steady state visually evoked potentials
TDD	true depth display
TIMIT	(unknown)
TOC	tactical operations center
UIAV	University of Illinois Active Vision System
UIUC	University of Illinois at Urbana-Champaign
VAT	visualization architecture technology
VGIS	virtual geographic information system
VLBV	very low bit rate video
VMD	(unknown)
VP	variable priority
VR	virtual reality
WWW	World Wide Web

INTENTIONALLY LEFT BLANK

1   ADMINISTRATOR
     DEFENSE TECHNICAL INFO CTR
     ATTN DTIC OCA
     8725 JOHN J KINGMAN RD STE 0944
     FT BELVOIR VA 22060-6218

1   DIRECTOR
     US ARMY RSCH LABORATORY
     ATTN AMSRL CI AI R REC MGMT
     2800 POWDER MILL RD
     ADELPHI MD 20783-1197

1   DIRECTOR
     US ARMY RSCH LABORATORY
     ATTN AMSRL CI LL TECH LIB
     2800 POWDER MILL RD
     ADELPHI MD 20783-1197

1   DIRECTOR
     US ARMY RSCH LABORATORY
     ATTN AMSRL D D SMITH
     2800 POWDER MILL RD
     ADELPHI MD 20783-1197

2   DIRECTOR
     US ARMY RSCH LABORATORY
     ATTN AMSRL CI J GANTT
           B FORNOFF
     2800 POWDER MILL RD
     ADELPHI MD 20783-1197

1   DIRECTOR
     US ARMY RSCH LABORATORY
     ATTN AMSRL CI C J GOWENS
     2800 POWDER MILL RD
     ADELPHI MD 20783-1197

1   DIRECTOR
     US ARMY RSCH LABORATORY
     ATTN AMSRL CI CN M HOLLAND
     2800 POWDER MILL RD
     ADELPHI MD 20783-1197

1   DIRECTOR
     US ARMY RSCH LABORATORY
     ATTN AMSRL D AD P J MILLER
     2800 POWDER MILL RD
     ADELPHI MD 20783-1197

1   BERNIE CORONA
     6263 AUDUBON DR
     COLUMBIA MD 21044-3814

1   ROCKWELL SCIENTIFIC COMPANY
     ATTN MARIUS VASSILIOU
     1049 CAMINO DOS RIO
     PO BOX 1085
     THOUSAND OAKS CA 91358-4633

1   MILESTONE GROUP
     ATTN JOHN F BROCK
     1600 WILSON BLVD STE 1200
     PO BOX 5586
     ARLINGTON VA 22209

1   DEPT OF COMPUTER SCIENCE
     ATTN OSCAR GARCIA
     1640 COLONEL GLENNE HWY
     DAYTON OH 45435-0001

1   DWB & ASSOCIATES
     ATTN WM O BLACKWOOD
     7205 RESERVOIR RD
     SPRINGFIELD VA 22150

2   NC AG & TECH STATE UNIV
     ATTN CELESTINE A NTUEN
           EUI H PARK
     1601 EAST MARKET ST
     GREENSBORO NC 27411 ·

1   BATTLE COMMAND BATTLE LAB
     ATTN MIKE FREEMAN
     FT LEAVENWORTH KS 66027-1344

1   BILL MARSHAK
     4433 DAYTON XENIA RD BLDG 1
     DAYTON OH 45432

1   USACECOM
     ATTN AMSEL RD C2 SS T
           LAKSHMI REBBAPRAGADA
     FT MONMOUTH NJ 07703

1   US MILITARY ACADEMY
     DEPT OF BEHAVIORAL SCIENCES &
           LEADERSHIP
     ATTN LTC LAWRENCE SHATTUCK
     WEST POINT NY 10996

1   DCSPER DIR OF MANPRINT
     ATTN DAPE MR
     300 ARMY PENTAGON
     WASHINGTON DC 20310-0300

NO. OF COPIES	ORGANIZATION	NO. OF COPIES	ORGANIZATION
1	BECKMAN INST UNIV OF ILLINOIS ATTN KATHIE ALBLINGER 405 N MATHEWS AVE URBANA IL 61801	1	DIRECTOR TDAD DCST ATTN ATTG C BLDG 161 FORT MONROE VA 23651-5000
1	DIR FOR PERS TECHNOLOGIES DPY CHIEF OF STAFF PERS 300 ARMY PENTAGON 2C733 WASHINGTON DC 20310-0300	1	HQ USAMRDC ATTN SGRD PLC FORT DETRICK MD 21701
1	DIR ARMY AUDIOLOGY & SPEECH CTR WALTER REED ARMY MED CTR WASHINGTON DC 20307-5001	1	CDR USA AEROMEDICAL RSCH LAB ATTN LIBRARY FORT RUCKER AL 36362-5292
1	OUSD(A)/DDDR&E(R&A)/E&LS PENTAGON ROOM 3D129 WASHINGTON DC 20301-3080	1	US ARMY SAFETY CTR ATTN CSSC SE FORT RUCKER AL 36362
1	CODE 1142PS OFC OF NAVAL RSCH 800 N QUINCY STREET ARLINGTON VA 22217-5000	1	CHIEF ARMY RSCH INST AVIATION R&D ACTIVITY ATTN PERI IR FORT RUCKER AL 36362-5354
1	WALTER REED INST OF RSCH ATTN SGRD UWI C COL REDMOND WASHINGTON DC 20307-5100	1	AIR FORCE FLIGHT DYNAMICS LAB ATTN AFWAL/FIES/SURVIAC WRIGHT PATTERSON AFB OH 45433
1	CDR US ARMY RSCH INST ATTN PERI ZT DR E M JOHNSON 5001 EISENHOWER AVENUE ALEXANDRIA VA 22333-5600	1	US ARMY NATICK RD&E CTR ATTN STRNC YBA NATICK MA 01760-5020
1	DEF LOGISTICS STUDIES INFORMATION EXCHANGE ATTN DIR DLSIE ATSZ DL BLDG 12500 2401 QUARTERS ROAD FORT LEE VA 23801-1705	1	US ARMY TROOP SUPPORT CMD NATICK RD&E CTR ATTN BEHAVIORAL SCI DIV SSD NATICK MA 01760-5020
1	HEADQUARTERS USATRADOC ATTN ATCD SP FORT MONROE VA 23651	1	US ARMY TROOP SUPPORT CMD NATICK RD&E CTR ATTN TECH LIB (STRNC MIL) NATICK MA 01760-5040
1	CDR USATRADOC COMMAND SAFETY OFC ATTN ATOS PESSAGNO/LYNE FORT MONROE VA 23651-5000	1	DR RICHARD JOHNSON HEALTH & PERFORMANCE DIV US ARIEM NATICK MA 01760-5007
		1	PROGRAM MANAGER RAH-66 ATTN SFAE AV RAH BLDG 5681 WOOD RD REDSTONE ARSENAL AL 35898

136

NO. OF
COPIES ORGANIZATION

1 NAVAL SUB MED RSCH LAB
MEDICAL LIB BLDG 148
BOX 900 SUBMARINE BASE
NEW LONDON
GROTON CT 06340

1 USAF ARMSTRONG LAB/CFTO
ATTN DR F W BAUMGARDNER
SUSTAINED OPERATIONS BR
BROOKS AFB TX 78235-5000

1 CDR
USAMC LOGISTICS SUP ACTIVITY
ATTN AMXLS AE
REDSTONE ARSENAL AL
35898-7466

1 ARI FIELD UNIT FT KNOX
BLDG 2423 PERI IK
FORT KNOX KY 40121-5620

1 CDR
WHITE SANDS MISSILE RANGE
ATTN STEWS TE RE
WSMR NM 88002

1 STRICOM
12350 RSCH PARKWAY
ORLANDO FL 32826-3276

1 PURDUE UNIVERSITY
SERIALS UNIT
CDM KARDEX
1535 STEWART CTR
WEST LAFAYETTE IN
47907-1535

1 GOVT PUBLICATIONS LIB
409 WILSON M
UNIVERSITY OF MINNESOTA
MINNEAPOLIS MN 55455

1 DR RICHARD PEW
BBN SYSTEMS &TECH CORP
10 MOULTON STREET
CAMBRIDGE MA 02138

1 DR HARVEY A TAUB
RSCH SECTION PSYCH SECTION
VETERANS ADMIN HOSPITAL
IRVING AVE & UNIVERSITY PLACE
SYRACUSE NY 13210

NO. OF
COPIES ORGANIZATION

1 DR ROBERT C SUGARMAN
132 SEABROOK DRIVE
BUFFALO NY 14221

1 DR ANTHONY DEBONS
IDIS UNIV OF PITTSBURGH
PITTSBURGH PA 15260

1 MR R BEGGS
BOEING-HELICOPTER CO
P30-18
PO BOX 16858
PHILADELPHIA PA 19142

1 DR ROBERT KENNEDY
ESSEX CORPORATION STE 227
1040 WOODCOCK ROAD
ORLANDO FL 32803

1 LAWRENCE C PERLMUTER PHD
UNIV OF HEALTH SCIENCES
THE CHICAGO MEDICAL SCHOOL
DEPT OF PSYCHOLOGY
3333 GREEN BAY ROAD
NORTH CHICAGO IL 60064

1 GENERAL DYNAMICS
LAND SYSTEMS DIV LIBRARY
PO BOX 1901
WARREN MI 48090

1 DR MM AYOUB DIRECTOR
INST FOR ERGONOMICS RSCH
TEXAS TECH UNIVERSITY
LUBBOCK TX 79409

1 DELCO DEF SYS OPERATIONS
ATTN RACHEL GONZALES B204
7410 HOLLISTER AVE
GOLETA CA 93117-2583

1 MR WALT TRUSZKOWSKI
NASA/GODDARD SPACE
FLIGHT CTR
CODE 588.0
GREENBELT MD 20771

NO. OF COPIES	ORGANIZATION	NO. OF COPIES	ORGANIZATION
1	US ARMY ATTN AVA GEDDES MS YA:219-1 MOFFETT FIELD CA 94035-1000	1	PROGRAM MANAGER RAH-66 ATTN SFAE AV RAH BLDG 5681 WOOD RD REDSTONE ARSENAL AL 35898
1	DR NORMAN BADLER DEPT OF COMPUTER & INFO SCIENCE UNIV OF PENNSYLVANIA PHILADELPHIA PA 19104-6389	1	JON TATRO HUMAN FACTORS SYS DESIGN BELL HELICOPTER TEXTRON INC PO BOX 482 MAIL STOP 6 FT WORTH TX 76101
1	CDR US ARMY RSCH INST OF ENVIRONMNTL MEDICINE NATICK MA 01760-5007	1	CHIEF CREW SYS INTEGRATION SIKORSKY AIRCRAFT M/S S3258 NORTH MAIN STREET STRATFORD CT 06602
1	HQDA (DAPE ZXO) ATTN DR FISCHL WASHINGTON DC 20310-0300	1	GENERAL ELECTRIC COMPANY ARMAMENT SYS DEPT RM 1309 ATTN HF/MANPRINT R C MCLANE LAKESIDE AVENUE BURLINGTON VT 05401-4985
1	HUMAN FACTORS ENG PROGRAM DEPT OF BIOMEDICAL ENGNG COLLEGE OF ENGINEERING & COMPUTER SCIENCE WRIGHT STATE UNIVERSITY DAYTON OH 45435	1	JOHN B SHAFER 250 MAIN STREET OWEGO NY 13827
		1	OASD (FM&P) WASHINGTON DC 20301-4000
1	CDR USA MEDICAL R&D COMMAND ATTN SGRD PLC LTC K FRIEDL FORT DETRICK MD 21701-5012	1	COMMANDANT US ARMY ARMOR SCHOOL ATTN ATSB CDS FT KNOX KY 40121-5215
1	PEO ARMOR SYS MODERNIATION US ARMY TANK-AUTOMOTIVE CMD ATTN SFAE ASM S WARREN MI 48397-5000	1	CDR US ARMY AVIATION CTR ATTN ATZQ CDM S FT RUCKER AL 36362-5163
1	PEO COMMUNICATIONS ATTN SFAE CM RE FT MONMOUTH NJ 07703-5000	1	CDR US ARMY SIGNAL CTR & FT GORDON ATTN ATZH CDM FT GORDON GA 30905-5090

1	PEO AIR DEF
	ATTN SFAE AD S
	US ARMY MISSILE COMMAND
	REDSTONE ARSENAL AL
	       35898-5750
		1	DIRECTOR
			US ARMY AEROFLIGHT
			    DYNAMICS DIR
			MAIL STOP 239-9
			NASA AMES RSCH CTR
			MOFFETT FIELD CA 94035-1000

1	PEO STRATEGIC DEF PO BOX 15280 ATTN DASD ZA US ARMY STRATEGIC DEF CMD ARLINGTON VA 22215-0280		

1   CDR
     MARINE CORPS SYSTEMS CMD
     ATTN   CBGT
     QUANTICO   VA 22134-5080

1   DIR AMC-FIELD ASSIST IN
          SCIENCE & TECHNOLOGY
     ATTN AMC-FAST
     FT BELVOIR   VA 22060-5606

1   CDR
     US ARMY FORCES CMD
     ATTN FCDJ SA   BLDG 600
     AMC FAST SCIENCE ADVISER
     FT MCPHERSON GA 30330-6000

1   CDR
     I CORPS AND FORT LEWIS
     AMC FAST SCIENCE ADVISER
     ATTN   AFZH CSS
     FORT LEWIS   WA 98433-5000

1   HQ III CORPS & FORT HOOD
     OFC OF THE SCIENCE ADVISER
     ATTN   AFZF CS SA
     FORT HOOD TX 76544-5056

1   CDR
     HQ XVIII ABN CORPS & FT BRAGG
     OFC OF THE SCI ADV BLDG 1-1621
     ATTN AFZA GD FAST
     FORT BRAGG   NC 28307-5000

1   SOUTHCOM WASHINGTON
               FIELD OFC
     1919 SOUTH EADS ST STE L09
     AMC FAST SCIENCE ADVISER
     ARLINGTON   VA 22202

1   HQ US SPECIAL OPERATIONS CMD
     AMC FAST SCIENCE ADVISER
     ATTN   SOSD
     MACDILL AIR FORCE BASE
     TAMPA   FL 33608-0442

1   HQ US ARMY EUROPE AND
               7TH ARMY
     ATTN   AEAGX SA
     OFC OF THE SCIENCE ADVISER
     APO AE 09014

1   CDR
     HEADQUARTERS USEUCOM
     AMC FAST SCIENCE ADVISER
     UNIT 30400 BOX 138
     APO AE 09128

1   HQ 7TH ARMY TRAINING CMD
     UNIT #28130
     AMC FAST SCIENCE ADVISER
     ATTN   AETT SA
     APO AE 09114

1   CDR
     HHC SOUTHERN EUROPEAN
               TASK FORCE
     ATTN AESE SA BLDG 98
     AMC FAST SCIENCE ADVISER
     APO AE 09630

1   CDR US ARMY PACIFIC
     AMC FAST SCIENCE ADVISER
     ATTN   APSA
     FT SHAFTER HI 96858-5L00

1   AMC FAST SCIENCE ADVISERS
     PCS #303 BOX 45 CS-SO
     APO AP 96204-0045

1   CDR & DIR USAE WATERWAYS
          EXPERIMENTAL STATION
     ATTN CEWES IM MI R
               A S CLARK
     CD DEPT #1153
     3909 HALLS FERRY ROAD
     VICKSBURG MS 39180-6199

1   ENGINEERING PSYCH LAB
     DEPT OF BEHAVIORAL
          SCIENCES & LEADERSHIP
     BLDG 601 ROOM 281
     US MILITARY ACADEMY
     WEST POINT NY 10996-1784

3   DIR SANDIA NATL LAB
     ENGNRNG MECHANICS DEPT
     MS 9042 ATTN J HANDROCK
          Y R KAN   J LAUFFER
     PO BOX 969
     LIVERMORE CA 94551-0969

140

NO. OF COPIES	ORGANIZATION
1	ARL HRED TACOM FLD ELMT ATTN AMSRL HR MU M SINGAPORE 6501 E 11 MILE RD MAIL STOP 284 BLDG 200A 2ND FL RM 2104 WARREN MI 48397-5000
1	ARL HRED USAFAS FLD ELMT ATTN AMSRL HR MF L PIERCE BLDG 3040 RM 220 FORT SILL OK 73503-5600
1	ARL HRED USAIC FLD ELMT ATTN AMSRL HR MW E REDDEN BLDG 4 ROOM 332 FT BENNING GA 31905-5400
1	ARL HRED USASOC FLD ELMT ATTN AMSRL HR MN R SPENCER DCSFDI HF HQ USASOC BLDG E2929 FORT BRAGG NC 28310-5000
1	ARL HRED HFID FLD ELMT ATTN AMSRL HR MP D UNGVARSKY BATTLE CMD BATTLE LAB 415 SHERMAN AVE UNIT 3 FT LEAVENWORTH KS 66027-2326
1	CDR AMC - FAST JRTC & FORT POLK ATTN AFZX GT DR J AINSWORTH CMD SCIENCE ADVISOR G3 FORT POLK LA 71459-5355
1	CDR HRED AMEDD ATTN AMSRL HR MM COL N VAUSE 2250 STANLEY RD STE 322 FT SAM HOUSTON TX 78234

ABERDEEN PROVING GROUND

NO. OF COPIES	ORGANIZATION
2	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRL CI LP (TECH LIB) BLDG 305 APG AA
1	LIBRARY ARL HRED BLDG 459

NO. OF COPIES	ORGANIZATION
1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRL CI N RADHAKRISHNAN BLDG 394
3	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRL HR SE P ROSE D HEADLEY L ALLENDER BLDG 459
1	ARL HRED ATTN AMSRL HR MB F PARAGALLO BLDG 459
1	ARL HRED ECBC FLD ELMT ATTN AMSRL HR M BLDG 459

INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE April 2002	3. REPORT TYPE AND DATES COVERED Final

4. TITLE AND SUBTITLE

Advanced Displays and Interactive Displays Report Compendium III—Final Report

5. FUNDING NUMBERS

AMS Code 622716
PR: AH70

6. AUTHOR(S)

Rose, Paul (ARL)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Research Laboratory
Human Research & Engineering Directorate
Aberdeen Proving Ground, MD 21005-5425

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Laboratory
Human Research & Engineering Directorate
Aberdeen Proving Ground, MD 21005-5425

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ARL-SR-115

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report contains more than 300 citations and abstracts of papers and presentations produced by the Advanced Displays and Interac-tive Displays consortium during the 5-year U.S. Army Federated Laboratory program. The program, more informally known as FedLab, was formed in 1996 to establish partnerships among the Army, industry, and academic research communities. The Advanced Displays and Interactive Displays consortium seeks to provide inno-vative, cost-effective solutions to information access, understanding, and management for the soldier of the future.

The research encompasses a range of topics. Some work concerns the representation of uncertainty and imprecision in databases or the representation of relationships in multimedia databases, in ways that are compatible with human cognitive-processing capabilities. Other work adopts the means of human communication (such as speech, gesture, eye gaze, and lipreading) for human-computer interaction. Additional work explores methods for incorporating information in virtual reality displays that support decision making without distracting or overwhelming the soldier. Although diverse, the research is linked by its overriding goal: the presentation of information in a form that allows effective human understanding and decision making in complex battlefield situations.

14. SUBJECT TERMS

abstracts
federated laboratories

15. NUMBER OF PAGES
150

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

143